

# Accelerating self-consistent field convergence with the augmented Roothaan–Hall energy function

Xiangqian Hu<sup>a)</sup> and Weitao Yang<sup>b)</sup>

*Department of Chemistry, Duke University, Durham, North Carolina 27708-0354, USA*

(Received 12 November 2009; accepted 12 January 2010; published online 4 February 2010)

Based on Pulay's direct inversion iterative subspace (DIIS) approach, we present a method to accelerate self-consistent field (SCF) convergence. In this method, the quadratic augmented Roothaan–Hall (ARH) energy function, proposed recently by Høst and co-workers [*J. Chem. Phys.* **129**, 124106 (2008)], is used as the object of minimization for obtaining the linear coefficients of Fock matrices within DIIS. This differs from the traditional DIIS of Pulay, which uses an object function derived from the commutator of the density and Fock matrices. Our results show that the present algorithm, abbreviated ADIIS, is more robust and efficient than the energy-DIIS (EDIIS) approach. In particular, several examples demonstrate that the combination of ADIIS and DIIS (“ADIIS+DIIS”) is highly reliable and efficient in accelerating SCF convergence.

© 2010 American Institute of Physics. [doi:10.1063/1.3304922]

## I. INTRODUCTION

Quantum mechanical (QM) methods play an important role in many research areas from elaborating the subtle reaction mechanisms of enzymes<sup>1,2</sup> to designing novel structures of molecules and materials.<sup>3–7</sup> Because of the good balance between computational cost and desired accuracy, Kohn–Sham (KS) density functional theory (KS-DFT),<sup>8,9</sup> based on a single Slater determinant, is a very popular method to describe the electronic structures of molecules or materials. In both Hatree–Fock and KS approaches, the self-consistent field (SCF) scheme is applied to obtain the total energies of systems. In this scheme, an initial guess for the density matrix ( $\mathbf{D}$ ) is employed to construct the Fock matrix (or KS matrix in DFT)  $[\mathbf{F}(\mathbf{D})]$ . Subsequently, an updated density matrix is solved and used to calculate the next Fock matrix, and this procedure is iterated until the density matrix becomes invariant (i.e., a SCF convergence is achieved). Unfortunately, SCF convergence without any accelerating technique is always problematic in many cases.<sup>10</sup> To achieve and accelerate the SCF convergence, a variety of methods has been developed such as simple density mixing between the previous and current density matrices, level shifting,<sup>11</sup> fractional electron occupations,<sup>12</sup> the optimal damping algorithm (ODA),<sup>13–15</sup> the direct inversion iterative subspace (DIIS) approach,<sup>16,17</sup> energy-DIIS (EDIIS),<sup>18,19</sup> density subspace minimization (DSM),<sup>20,21</sup> and the ground-state-directed optimization scheme.<sup>22,23</sup> Since the DIIS-based algorithms are particularly robust and efficient in most molecular systems, we focus on the DIIS procedure for SCF convergence in this work.

As emphasized in Ref. 23, the standard SCF iterative procedure with the DIIS-based methods involves two separate steps. The first step is to diagonalize the Fock matrix in order to construct a new density matrix; the second step is to

improve the new density matrix using the DIIS scheme to combine linearly this new density matrix with the density matrices obtained from previous iterations. In the standard DIIS approach developed by Pulay,<sup>16,17</sup> the linear coefficients of each density matrix are optimized by minimizing the orbital rotation gradient based on the commutator matrix of the Fock and density matrices (i.e.,  $[\mathbf{F}(\mathbf{D}), \mathbf{D}]$ ) in the orthonormal basis space. However, the minimization of this orbital rotation gradient does not always lead to a lower energy, particularly when the SCF is not close to convergence.<sup>18</sup> This may cause large energy oscillations and divergence in the SCF procedure. To stabilize DIIS, the EDIIS approach developed by Scuseria and co-workers<sup>18</sup> minimizes a quadratic energy function derived from the ODA<sup>13–15</sup> to obtain the linear coefficients in DIIS. This energy minimization driven EDIIS method rapidly brings the density matrix from the initial guess to a convergent region. In addition, the combination of EDIIS and DIIS makes SCF robust and efficient. Although “EDIIS+DIIS” is successful in accelerating the SCF convergence for many cases, the quadratic energy function used in EDIIS is only accurate for HF. In KS-DFT, approximate quadratic interpolation of energies<sup>18</sup> is employed because of the nonlinearity of the exchange-correlation functional. Therefore, the reliability of EDIIS can be impaired by the interpolation accuracy.

Here, we present a new approach based on the augmented Roothaan–Hall (ARH) energy function<sup>22,23</sup> to obtain the linear coefficients for the density matrices in DIIS. We first illustrate the mathematical expressions defining the ARH energy in Sec. II A. Then, we discuss the combination of the ARH energy and DIIS (ADIIS in short) and compare ADIIS with the standard DIIS and EDIIS algorithms in Sec. II B. We performed calculations on several molecular systems to validate the efficiency of ADIIS in comparison with DIIS and EDIIS, and report these results in Sec. III. Finally, we summarize our work in Sec. IV.

<sup>a)</sup>Electronic mail: xqhu@duke.edu.

<sup>b)</sup>Electronic addresses: weitao.yang@duke.edu and yang@chem.duke.edu.

## II. METHODS

### A. The ARH energy function

In Ref. 22, Høst and co-workers presented a quadratic ARH energy function to optimize directly the density matrix in the SCF procedure. Here, the key points about the ARH energy function for a closed-shell system are explained. (For an open-shell system, the  $\alpha$  and  $\beta$  spins need to be considered separately.) The total HF or KS-DFT energy for a closed-shell system can be expanded to second order, with respect to the density matrix, using the Taylor expansion,

$$E(\mathbf{D}) \approx \tilde{E}(\mathbf{D}) = E(\mathbf{D}_n) + \langle \mathbf{D} - \mathbf{D}_n | E^{[1]}(\mathbf{D}_n) \rangle + \frac{1}{2} \langle \mathbf{D} - \mathbf{D}_n | E^{[2]}(\mathbf{D}_n) | \mathbf{D} - \mathbf{D}_n \rangle, \quad (1)$$

where  $E(\mathbf{D})$  is the total energy of a density matrix  $\mathbf{D}$ ,  $\tilde{E}(\mathbf{D})$  is the approximate ARH energy function,  $\mathbf{D}_n$  is the density matrix of the  $n$ th SCF iteration, and  $\langle \mathbf{A} | \mathbf{B} \rangle = \text{Tr}(\mathbf{A}^T \mathbf{B})$ . In Eq. (1), the first derivative of  $E(\mathbf{D}_n)$  [i.e.,  $E^{[1]}(\mathbf{D}_n)$ ] with respect to  $\mathbf{D}_n$  is the corresponding Fock matrix [i.e.,  $\mathbf{F}(\mathbf{D}_n)$ ],

$$E^{[1]}(\mathbf{D}_n) = \left. \frac{\partial E(\mathbf{D})}{\partial \mathbf{D}} \right|_{\mathbf{D}=\mathbf{D}_n} = 2\mathbf{F}(\mathbf{D}_n) = 2\mathbf{F}_n. \quad (2)$$

The second derivative of  $E(\mathbf{D}_n)$  with respect to  $\mathbf{D}_n$  [i.e.,  $E^{[2]}(\mathbf{D}_n)$ ] is much more computationally expensive to evaluate exactly from the analytical expressions. A useful quasi-Newton approximation to  $E^{[2]}(\mathbf{D}_n)$  is imposed to circumvent this costly calculation,<sup>22</sup>

$$E^{[2]}(\mathbf{D}_n)(\mathbf{D} - \mathbf{D}_n) \approx E^{[1]}(\mathbf{D}) - E^{[1]}(\mathbf{D}_n) = 2\mathbf{F}(\mathbf{D}) - 2\mathbf{F}(\mathbf{D}_n). \quad (3)$$

Thus Eq. (1) for a closed-shell system is rewritten as

$$E(\mathbf{D}) \approx E(\mathbf{D}_n) + 2\langle \mathbf{D} - \mathbf{D}_n | \mathbf{F}(\mathbf{D}_n) \rangle + \langle \mathbf{D} - \mathbf{D}_n | [\mathbf{F}(\mathbf{D}) - \mathbf{F}(\mathbf{D}_n)] \rangle. \quad (4)$$

Since the density matrix must satisfy the symmetry ( $\mathbf{D} = \mathbf{D}^T$ ), trace [ $\text{Tr}(\mathbf{D}\mathbf{S}) = N_{\text{elec}}$ , where  $\mathbf{S}$  is the overlap matrix of atomic basis sets and  $N_{\text{elec}}$  is the total number of electrons], and idempotency ( $\mathbf{D}\mathbf{S}\mathbf{D} = \mathbf{D}$ ) constraints, the elements of  $\mathbf{D}$  cannot be optimized directly in Eq. (4). In Ref. 22, an anti-symmetric matrix is utilized to parametrize the elements of  $\mathbf{D}$  and fulfill all three constraints on the density matrix. Combined with the trust-region method,<sup>24</sup> direct optimization of the density matrix without diagonalization was carried out. Their approach was demonstrated to be robust and efficient, particularly when the classical DIIS approaches with diagonalization fail to converge, or converge to a saddle point instead of a minimum.

However, the diagonalization step in DIIS is the most efficient method to obtain the coefficients and eigenvalues of molecular orbitals as well as the density matrix. In particular, the diagonalization cost is low for many interesting moderately sized systems, such as enzymatic active sites (treated using HF or KS-DFT) in QM/MM simulations.<sup>1</sup> Therefore, in this work, we propose a new method combining the ARH energy function and the standard DIIS approach, abbreviated as ADIIS, to improve the efficiency and reliability of the SCF convergence.

### B. The ADIIS algorithm

Let us assume that  $\mathbf{D}_1, \mathbf{D}_2, \dots$ , and  $\mathbf{D}_n$  are  $n$  density matrices and  $\mathbf{F}_1, \mathbf{F}_2, \dots$ , and  $\mathbf{F}_n$  are the corresponding Fock matrices after  $n$  SCF iterations are performed. Similar to EDIIS,<sup>18</sup> the approximate density matrix  $\tilde{\mathbf{D}}_{n+1}$  for the  $(n+1)$ th iteration can be linearly expanded in terms of the previous density matrices and the linear coefficients can be obtained readily and efficiently via energy minimization

$$\tilde{\mathbf{D}}_{n+1} = \arg \min \left\{ E(\tilde{\mathbf{D}}), \tilde{\mathbf{D}} = \sum_{i=1}^n c_i \mathbf{D}_i, \sum_{i=1}^n c_i = 1, c_i \geq 0 \right\}. \quad (5)$$

Here, a convex linear combination of density matrices (i.e.,  $c_i \in [0, 1]$ )<sup>18</sup> as in EDIIS is imposed to simplify and stabilize the optimization. Once  $\{c_i, i=1, \dots, n\}$  are obtained via energy minimization, the Fock matrix  $\tilde{\mathbf{F}}_{n+1}$  can be evaluated by Pulay's DIIS scheme

$$\tilde{\mathbf{F}}_{n+1} = \sum_{i=1}^n c_i \mathbf{F}_i. \quad (6)$$

Then the new density matrix  $\mathbf{D}_{n+1}$ , fulfilling the symmetry, trace, and idempotency constraints, can be constructed by diagonalization of  $\tilde{\mathbf{F}}_{n+1}$ . In this work, to obtain the linear coefficients in Eq. (5), the ARH energy function [defined by Eq. (4)] is used to compute  $E(\tilde{\mathbf{D}})$ , which leads to the ARH-energy DIIS or ADIIS scheme. Using the notation  $f^{\text{ADIIS}} = \tilde{E}(\sum_{i=1}^n c_i \mathbf{D}_i)$ , Eq. (4) can be rewritten as

$$f^{\text{ADIIS}}(c_1, \dots, c_n) = E(\mathbf{D}_n) + 2 \sum_{i=1}^n c_i \langle \mathbf{D}_i - \mathbf{D}_n | \mathbf{F}(\mathbf{D}_n) \rangle + \sum_{i=1}^n \sum_{j=1}^n c_i c_j \langle \mathbf{D}_i - \mathbf{D}_n | [\mathbf{F}(\mathbf{D}_j) - \mathbf{F}(\mathbf{D}_n)] \rangle. \quad (7)$$

In contrast, for a closed-shell system, the energy expression  $E(\tilde{\mathbf{D}})$  used in EDIIS is given by

$$f^{\text{EDIIS}}(c_1, \dots, c_n) = \sum_{i=1}^n c_i E(\mathbf{D}_i) - \sum_{i=1}^n \sum_{j=1}^n c_i c_j \langle \mathbf{D}_i - \mathbf{D}_j | \mathbf{F}_i - \mathbf{F}_j \rangle. \quad (8)$$

ADIIS is based on the second order Taylor expansion of the total energy with respect to the density matrix and hence  $f^{\text{ADIIS}}$  is accurate for HF and KS-DFT calculations only if the quasi-Newton condition is sufficient. In contrast,  $f^{\text{EDIIS}}$  is only precise for HF calculations because the HF energy is quadratic in the density matrix. But  $f^{\text{EDIIS}}$  is an approximate quadratic expression for KS-DFT because the exchange-correlation term in KS-DFT is nonlinear in the density matrix.<sup>18</sup> (Note that the DSM-based method<sup>20,21</sup> only uses up to the first order Taylor expansion as the energy function to optimize linear coefficients  $\{c_i\}$  in the purified density matrix.)

For EDIIS or ADIIS, the optimization problem is

$$\min \left\{ f^{\text{EDIIS}} \text{ or } f^{\text{ADIIS}}, \sum_{i=1}^n c_i = 1, c_i \geq 0 \right\}. \quad (9)$$

For DIIS, the optimization problem is

$$\min \left\{ f^{\text{DIIS}}, \sum_{i=1}^n c_i = 1 \right\}, \quad (10)$$

where  $f^{\text{DIIS}}$  is defined as

$$f^{\text{DIIS}}(c_1, \dots, c_n) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j [\mathbf{F}_i, \mathbf{D}_i \mathbf{S}] \cdot [\mathbf{F}_j, \mathbf{D}_j \mathbf{S}]. \quad (11)$$

Based on Eqs. (9) and (10), ADIIS, similar to EDIIS, is indeed an *interpolation* scheme with the convex constraints  $\{c_i \in [0, 1]\}$ . Pulay's standard DIIS is an *extrapolation* procedure without the convex constraints. Moreover, ADIIS utilizes a clearly defined energy expression [Eq. (1)] to optimize the linear coefficients  $\{c_i\}$  toward the energy minimum, while DIIS only minimizes the orbital rotation gradient, which does not necessarily lead toward the energy minimum.<sup>18</sup>

As explained in Ref. 18 by Kudin *et al.*, the energy functions in both ADIIS and EDIIS are less sensitive than the orbital rotation gradient in DIIS when the density matrix is varied. As such, ADIIS and EDIIS may be less efficient than DIIS in the region close to SCF convergence. Therefore, "ADIIS+DIIS," similar to EDIIS+DIIS in Ref. 18, was also implemented in this work. In ADIIS+EDIIS, ADIIS is carried out at the beginning of the SCF to obtain quickly a nearly converged density matrix and then ADIIS is switched over to DIIS to achieve rapid final convergence.

### III. RESULTS

All the algorithms, including DIIS, EDIIS, and ADIIS, were implemented in an in-house program.<sup>25</sup> Several molecules were chosen to illustrate the performance of ADIIS and ADIIS+DIIS in comparison with DIIS, EDIIS, and EDIIS+DIIS. These are CH<sub>3</sub>CHO,<sup>18</sup> a water cluster,<sup>22</sup> a cadmium complex,<sup>22</sup> a polyaniline peptide,<sup>22</sup> and two ruthenium compounds. (All of the geometry structures are available via <http://www.chem.duke.edu/~xqhu/geoms.tar.gz>) In the standard DIIS calculations, neither level shifting or density damping is added. For CH<sub>3</sub>CHO and the cadmium complex, the core Hamiltonians were used to obtain the initial density matrix. For the water cluster, the polyaniline peptide, and two Ruthenium compounds, the atomic density matrices were employed to construct the initial density matrix. SCF convergence is considered being achieved when the energy difference between the current iteration and the previous one is less than 10<sup>-8</sup> a.u. For ADIIS+DIIS and EDIIS+DIIS, we used a simple criterion to activate DIIS: when the energy difference between two SCF cycles is less than 0.01 a.u. (2.0 a.u. was used for the polyaniline peptide and 1.0 a.u. was used for two Ruthenium compounds), ADIIS or EDIIS is switched to DIIS. (Note that other criteria<sup>18</sup> can be also applied to activate DIIS.) To obtain the linear coefficients for EDIIS and ADIIS, the L-BFGS algorithm<sup>26,27</sup> was used and the energy functions [Eq. (8) for EDIIS and Eq. (7) for

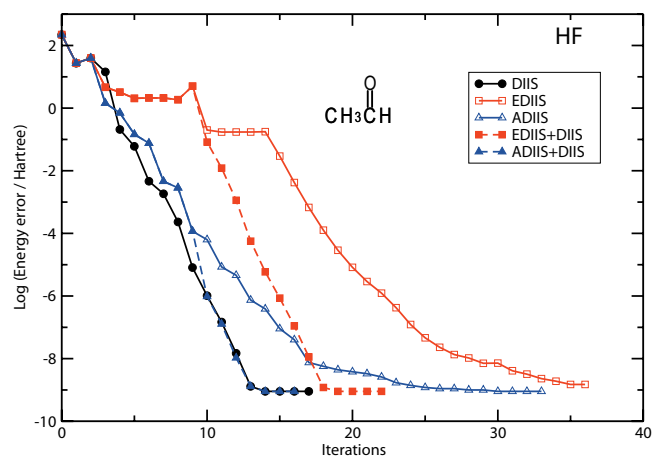


FIG. 1. Convergence of the CH<sub>3</sub>CHO energy with DIIS (the black solid line with circles), EDIIS (the red solid line with empty squares), ADIIS (the blue solid line with empty triangles), EDIIS+DIIS (the red dashed line with squares), and ADIIS+DIIS (the blue dashed line with triangles). A core Hamiltonian was applied to obtain the initial density matrix. The total energy was calculated with HF/6-31G\*. The converged energy is -152.914 4029 a.u.

ADIIS] were optimized using Eq. (9) with the aid of first derivatives ( $\partial f^{\text{EDIIS}}/\partial c_i$  and  $\partial f^{\text{ADIIS}}/\partial c_i$ , respectively). [Note that variable substitutions ( $c_i = t_i^2 / \sum_j t_j^2$ ) are applied to transform the constrained optimization problem into an unconstrained one in L-BFGS.] The averaged coefficients were taken as the initial guess during the L-BFGS optimizations. In this work, only six Fock matrices were combined linearly to construct  $\tilde{\mathbf{F}}_{n+1}$  in Eq. (6).

#### A. Well-behaved molecular systems: CH<sub>3</sub>CHO and a water cluster

We first chose CH<sub>3</sub>CHO (acetaldehyde) as a simple and well-behaved small molecule to illustrate the performance of DIIS, EDIIS, ADIIS, EDIIS+DIIS, and ADIIS+DIIS. Calculations were performed at the RHF/6-31G\* (Refs. 28 and 29) level of theory (53 Gaussian basis functions). As shown in Fig. 1, compared to the standard DIIS approach, both EDIIS and ADIIS are much slower to reach SCF convergence because both EDIIS and ADIIS are based on minimizing the approximate energy functions [Eq. (8) and Eq. (7)] at each SCF cycle. The energy functions are not sensitive to small variations in density matrices, particularly when the density matrix is close to the convergence region. In contrast, the orbital rotation gradient minimized in DIIS is more sensitive to the variations of density matrices. Overall, DIIS is much faster than ADIIS and EDIIS to converge the energy of simple small organic molecules, which is consistent with the discussions in Ref. 18. However, in the first few SCF iterations, EDIIS and ADIIS are capable of getting a lower energy from the crude initial guess than the DIIS energy (for instance, see the energies at the fourth cycle for EDIIS and ADIIS). Therefore, similar to the EDIIS+DIIS approach,<sup>18</sup> we also combined ADIIS with DIIS (i.e., ADIIS+DIIS): in the beginning of the SCF, ADIIS is applied to accelerate the SCF procedure and is switched over to DIIS when the energy difference between two SCF cycles is less than the specified threshold (0.01 a.u. is the default value in this work.). As

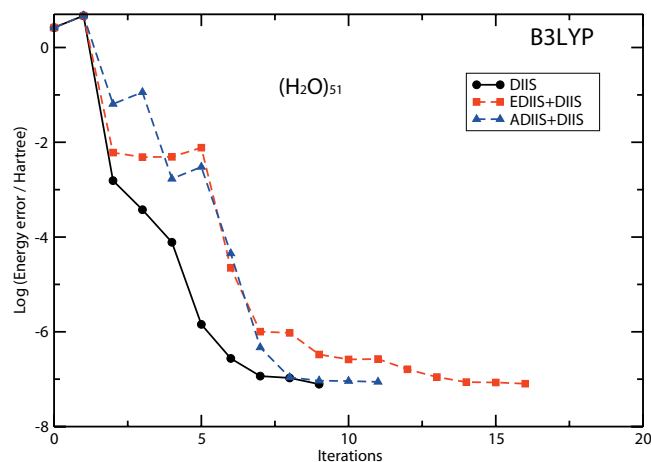


FIG. 2. Convergence of the energy for a water cluster consisting of 51 monomers with DIIS (the black solid line with circles), EDIIS+DIIS (the red dashed line with squares), and ADIIS+DIIS (the blue dashed line with triangles). The initial density matrix was constructed from the atomic density matrices. The total energy was calculated with B3LYP/6-31G\*\*. The converged energy is  $-3898.232\ 5683$  a.u.

such, ADIIS helps to obtain a superior density matrix close to the convergence region from the initial guess in the beginning of the SCF and makes DIIS more reliable and robust afterward. For  $\text{CH}_3\text{CHO}$ , when DIIS was switched on within ten SCF cycles, ADIIS+DIIS and EDIIS+DIIS significantly improved the speed of SCF convergence, as shown in Fig. 1. In addition, the efficiency of ADIIS+DIIS was the same as DIIS in this case. Since ADIIS+DIIS or EDIIS+DIIS is much more efficient than ADIIS or EDIIS alone, we only compared these two methods with DIIS for the other molecules. A much bigger system, a water cluster with 51 monomers<sup>22</sup> using B3LYP(VWN5)<sup>30-33</sup>/6-31G\* (1275 Gaussian basis functions), was also considered and the results were shown in Fig. 2. Since a good initial guess for density matrix was obtained from the atomic density matrices, DIIS was faster than ADIIS+DIIS and EDIIS+DIIS in terms of the number of SCF cycles required for convergence. Hence, for well-behaved systems such as  $\text{CH}_3\text{CHO}$  and the water cluster, DIIS is the most efficient algorithm for accelerating SCF convergence.<sup>18</sup> The efficiency of our ADIIS+DIIS approach is comparable to DIIS and is more efficient than EDIIS+DIIS in such well-behaved systems.

### B. A challenging system: A cadmium complex

A  $\text{Cd}^{2+}$ -imidazole complex<sup>22</sup> at both the RHF and B3LYP(VWN5)/3-21G<sup>34,35</sup> levels (92 Gaussian basis functions) is a challenging system. For this cadmium-imidazole compound, the standard DIIS scheme failed to converge for HF and B3LYP (see Fig. 3) when the core Hamiltonian was employed to obtain the initial density matrix. However, EDIIS+DIIS and ADIIS+DIIS can achieve SCF convergence efficiently with HF, as shown in Fig. 3(a). For DFT methods such as B3LYP, EDIIS+DIIS needed almost 60 SCF iterations to accumulate enough information to bring the density matrix close to convergence. These different convergence behaviors of EDIIS+DIIS are due to the energy expression in EDIIS, which is precise for HF and approxi-

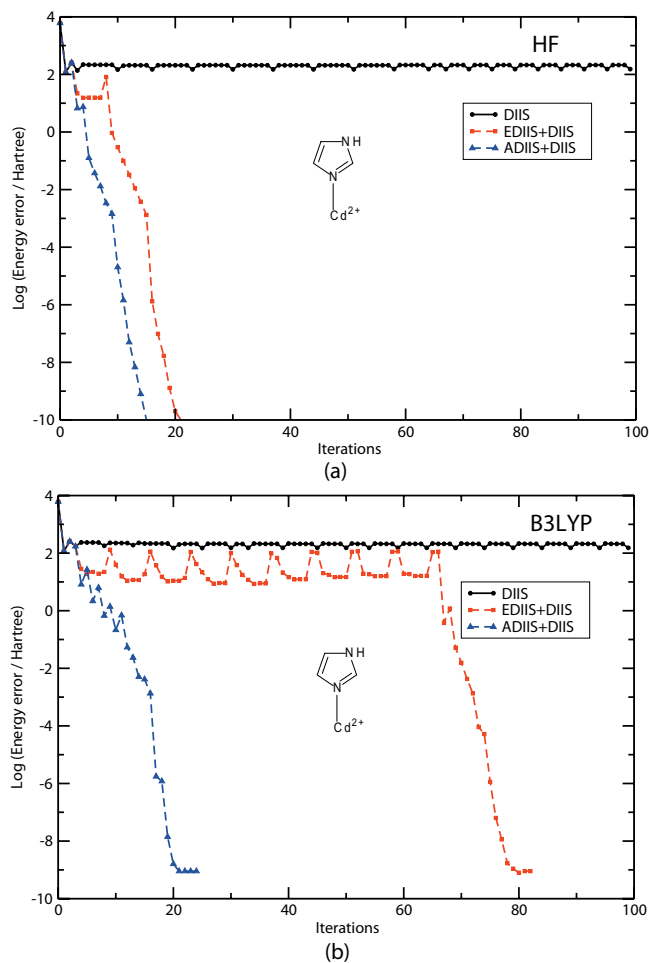


FIG. 3. Convergence of the energy for a cadmium-imidazole complex with DIIS (the black solid line with circles), EDIIS+DIIS (the red dashed line with squares), and ADIIS+DIIS (the blue dashed line with triangles). A core Hamiltonian was applied to obtain the initial density matrix. The total energy was calculated with RHF/3-21G (a) and with B3LYP/3-21G (b). The converged energies are  $-5663.143\ 3914$  a.u. for RHF and  $-5666.922\ 5744$  a.u. for B3LYP.

mately calculated by the quadratic ODA for DFT. The ODA scheme can be problematic for some molecules with DFT. Compared to DIIS and EDIIS+DIIS, ADIIS+DIIS minimized the energy robustly in the beginning of the SCF via the ADIIS scheme and it proceeded through the SCF iterations smoothly and efficiently after DIIS was switched on in both the HF and DFT cases [see Figs. 3(a) and 3(b)]. This suggests that the ARH energy function [i.e., Eq. (7)] is good approximation both for HF and DFT. The Taylor expansion up to second order with respect to the density matrix is sufficiently accurate to predict the next improved Fock matrix  $\tilde{\mathbf{F}}_{n+1}$  in Eq. (6). Hence, during the SCF procedure, the density matrix of the cadmium-imidazole system was rapidly optimized from the crude initial guess by ADIIS for HF or DFT, which significantly improved the efficiency and reliability of the subsequent DIIS in ADIIS+DIIS.

### C. A typical biological system: A polyaniline peptide

To further illustrate the performance of DIIS, EDIIS+DIIS, and ADIIS+DIIS in a normal biological system, energy calculations were carried out for a polyaniline peptide



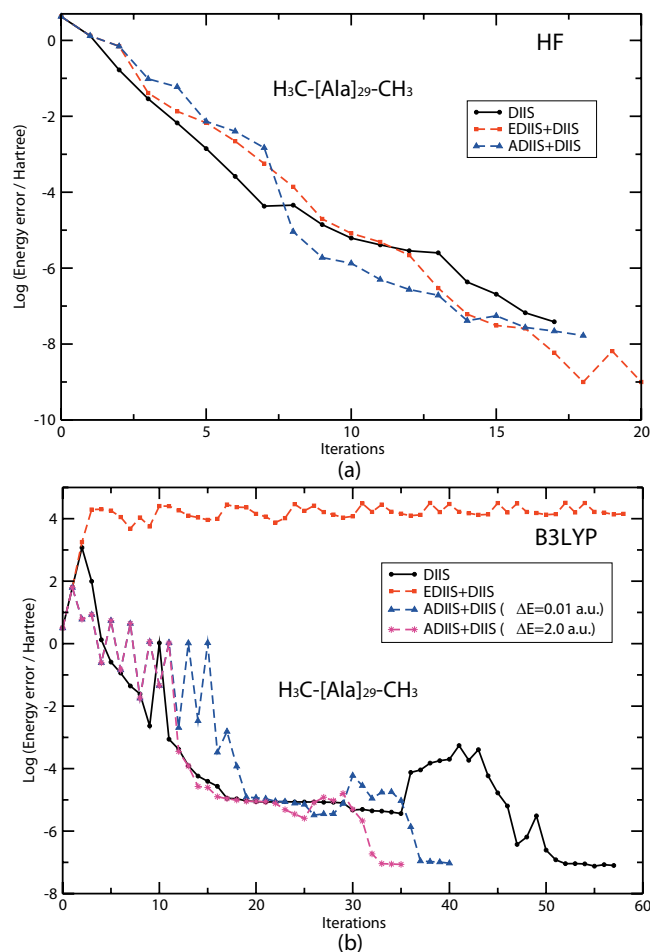


FIG. 4. Convergence of the energy of a 29-residue polyaniline peptide with DIIS (the black solid line with circles), EDIIS+DIIS (the red dashed line with squares), ADIIS+DIIS with the energy difference threshold of 0.01 a.u. to activate DIIS (the blue dashed line with triangles), and ADIIS+DIIS with the threshold of 2.0 a.u. (the magenta dashed line with stars). The initial density matrix was constructed from the atomic density matrices. The total energy was calculated by RHF/6-31G and B3LYP/6-31G. The converged energies are  $-7127.568\ 1797$  a.u. for RHF and  $-7171.566\ 7248$  a.u. for B3LYP.

with 29 residues<sup>22</sup> with RHF and B3LYP(VWN5)/6-31G<sup>36,37</sup> (1599 Gaussian basis functions). As shown in Fig. 4(a), for all of three algorithms, the energy decreased monotonically through the SCF iterations with HF and convergence was achieved within 20 cycles. In contrast, for the B3LYP results in Fig. 4(b), EDIIS+DIIS failed to converge because the energy error in Eq. (8) from the exchange-correlation energy in DFT is larger due to error accumulation as the chemical system size becomes larger. In the first few SCF iterations, the energy obtained with EDIIS+DIIS actually increased and the density matrix was trapped in the incorrect electronic state. For DIIS and ADIIS+DIIS, both algorithms eventually converged to the energy minimum. However, DIIS first converged to a saddle point and required more than 15 SCF cycles to jump out this critical point [see the broad peak in Fig. 4(b) for DIIS]. Compared to the total 57 SCF cycles for DIIS, ADIIS+DIIS needed fewer cycles to overcome the saddle point, and it reached the energy minimum within 40 SCF iterations. When a larger energy threshold was used ( $\Delta E=2.0$  a.u.) to switch on DIIS, ADIIS+DIIS also

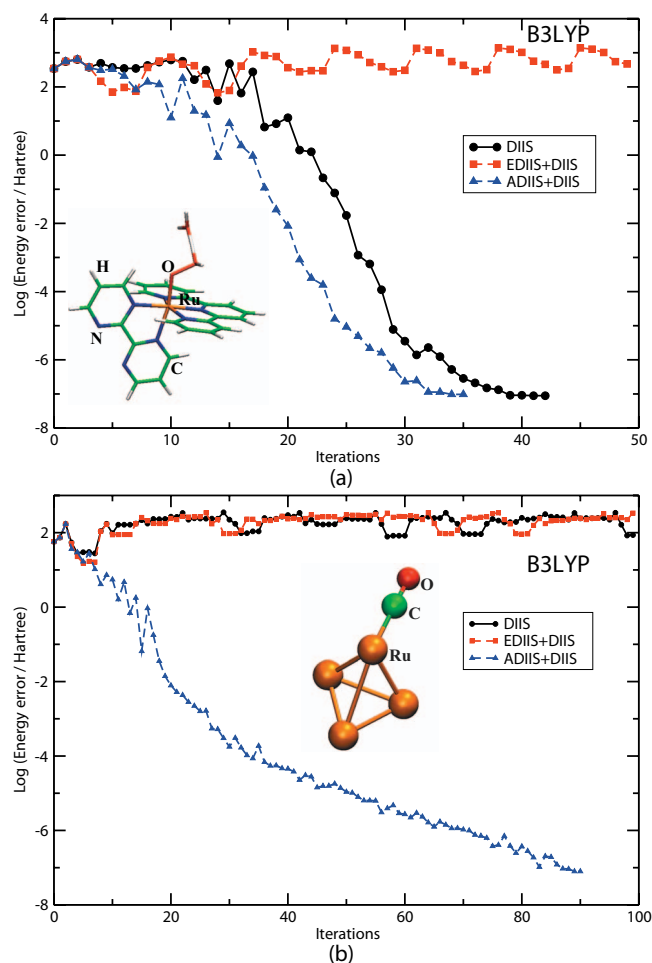


FIG. 5. Convergence of the energies of  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  and  $\text{Ru}_4(\text{CO})$  with DIIS (the black solid line with circles), EDIIS+DIIS (the red dashed line with squares), and ADIIS+DIIS with the energy difference threshold of 1.0 a.u. to activate DIIS (the blue dashed line with triangles). The initial density matrix was constructed from the atomic density matrices. The total energy was calculated by B3LYP/LanL2dz.  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  has 123 $\alpha$  spin electrons and 122.5 $\beta$  spin electrons.  $\text{Ru}_4(\text{CO})$  is closed shell. The converged energies are  $-1595.849\ 0175$  a.u. for  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  and  $-488.710\ 6840$  a.u. for  $\text{Ru}_4(\text{CO})$ .

achieved SCF convergence. This indicates that the performance of ADIIS+DIIS does not depend on the energy threshold critically.

#### D. Problem cases: Two Ruthenium compounds

We further studied the SCF convergence behavior for two ruthenium compounds:  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  and  $\text{Ru}_4(\text{CO})$  with B3LYP(VWN5)/LanL2dz<sup>38</sup> and the results are shown in Fig. 5.  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  is  $[\text{Ru}(\text{tpy})(\text{bpm})(\text{OH}_2)]^{2.5+}$  (tpy is 2,2':6',2''-terpyridine; bpm is 2,2'-bipyrimidine) that is an open-shell system with a fractional electron, 363 Gaussian basis functions, and 56 157 TIP3P water molecules.<sup>39</sup>  $[\text{Ru}^{\text{V}}=\text{O}]^{+2.5}$  has 123 $\alpha$  spin electrons and 122.5 $\beta$  spin electrons with hundreds of thousands of TIP3P water molecules in the system. The fractional electronic occupation has been successfully applied in redox potential simulations<sup>40</sup> and it is very useful to study the quality of approximate density functionals in DFT.<sup>41-43</sup> However, many systems with the fractional electronic occupations often suffer SCF convergence

problems. As shown in Fig. 5(a), EDIIS+DIIS failed to achieve the convergence. Our ADIIS+DIIS method is more stable and robust than DIIS for such a fractional-charge open-shell system.

For the second example,  $\text{Ru}_4(\text{CO})$  is a closed-shell system with 114 Gaussian basis functions and its geometry was taken from Ref. 44. This compound was used to study CO absorption on transition metal clusters.<sup>44</sup> Fig. 5(b) shows that both EDIIS+DIIS and DIIS failed to achieve SCF convergence for  $\text{Ru}_4(\text{CO})$ . However, for ADIIS+DIIS, the total energy of  $\text{Ru}_4(\text{CO})$  decreased monotonically, reaching convergence after 90 iterations. Thus, ADIIS+DIIS seems to be more effective and robust than EDIIS+DIIS and DIIS for heavy transition metal systems in both open- and closed-shell cases.

In summary, various molecular systems, ranging from simple organic molecules to biomolecules, from closed-shell to open-shell cases, and from light atoms to transition metals, were computed by several SCF algorithms. ADIIS+DIIS with both HF and DFT is efficient enough to guide the construction of updated density matrices in the SCF process for the systems studied. The density matrices explored by ADIIS in the beginning stages of the SCF make the standard DIIS more robust and reliable afterward. ADIIS+DIIS provides a robust and efficient protocol to accelerate SCF convergence.

#### IV. CONCLUSIONS

In the standard DIIS approach, the orbital rotation gradient is minimized to obtain the linear Fock matrix coefficients, then a new Fock matrix is constructed for diagonalization in the SCF. However, the gradient minimum may not correspond to the correct energy minimum, which causes energy oscillations during the SCF [for instance, see Figs. 3 and 5(b)]. Although EDIIS minimizes the energy function derived from the ODA in each SCF cycle, the energy expression is only accurate for HF. For DFT, EDIIS can have large errors and fail to achieve SCF convergence in some molecular systems [for instance, see Fig. 4(b)]. In this work, we applied a more rigorously defined ARH energy function to optimize the linear coefficients of the Fock matrices using BFGS. This newly developed ADIIS scheme rapidly brings the initial density matrix close to the convergence region. Several examples in this paper demonstrate that after ADIIS is combined with DIIS to yield ADIIS+DIIS, it accelerates SCF convergence more reliably and efficiently than DIIS, EDIIS, and EDIIS+DIIS.

It is worth noting that the quadratic ARH energy function combined with trust-region optimization was originally developed to accelerate SCF convergence without Fock matrix diagonalization.<sup>22,23</sup> In those studies, two separate computational steps (i.e., diagonalization and DIIS averaging) were no longer involved in each SCF cycle and direct optimization of the density matrix was performed. However, the diagonalization step can be considered as a direct minimization of a quadratic model of the total energy and is still a most efficient scheme to obtain useful molecular orbital coefficients and the corresponding density matrix in each SCF cycle. Particularly, the diagonalization step is not too costly

for many interesting molecular systems (even for systems with thousands of basis functions). To capitalize on the efficiency from diagonalization, the new ADIIS method makes use of the ARH energy function derived from the Taylor expansion to optimize the linear Fock matrix coefficients, while the Fock matrices are based on density matrices constructed from diagonalization in previous SCF steps. Furthermore, ADIIS+DIIS can be easily coded in current quantum computational packages. Our results indicate that ADIIS dramatically improves the quality of density matrix at the beginning of the SCF and helps DIIS converge to the minimum rapidly in ADIIS+DIIS.

In addition to the acceleration techniques, the exchange-correlation functional form in DFT could also affect the SCF convergence behavior. In particular, the intrinsic delocalization and static correlation errors<sup>41-43</sup> in HF and approximate DFT functionals can force the SCF to converge to a unphysical electronic state. We believe that the erratic or divergent behavior of the SCF procedure can be relieved by the next generation of DFT functionals with minimal delocalization and static correlation errors. More importantly, ADIIS is extendable to any new functional form because the ARH energy function is only based on the second order Taylor expansion with respect to the density matrix, with the Hessians approximated by a quasi-Newton approach.

#### ACKNOWLEDGMENTS

We acknowledge the support from the National Institutes of Health and the National Science Foundation. Discussions with Dr. Erin R. Johnson have been helpful.

- <sup>1</sup>H. Hu and W. T. Yang, *Annu. Rev. Phys. Chem.* **59**, 573 (2008).
- <sup>2</sup>D. J. Tantillo, J. G. Chen, and K. N. Houk, *Curr. Opin. Chem. Biol.* **2**, 743 (1998).
- <sup>3</sup>M. L. Wang, X. Q. Hu, D. N. Beratan, and W. T. Yang, *J. Am. Chem. Soc.* **128**, 3228 (2006).
- <sup>4</sup>S. Keinan, X. Q. Hu, D. N. Beratan, and W. T. Yang, *J. Phys. Chem. A* **111**, 176 (2007).
- <sup>5</sup>X. Q. Hu, D. N. Beratan, and W. T. Yang, *J. Chem. Phys.* **129**, 9 (2008).
- <sup>6</sup>W. L. Jorgensen, *Science* **303**, 1813 (2004).
- <sup>7</sup>A. Franceschetti and A. Zunger, *Nature (London)* **402**, 60 (1999).
- <sup>8</sup>R. G. Parr and W. T. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- <sup>9</sup>W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- <sup>10</sup>W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, Weinheim, 2002).
- <sup>11</sup>V. R. Saunders and I. H. Hillier, *Int. J. Quantum Chem.* **7**, 699 (1973).
- <sup>12</sup>A. D. Rabuck and G. E. Scuseria, *J. Chem. Phys.* **110**, 695 (1999).
- <sup>13</sup>E. Cancès and C. Le Bris, *Math. Modell. Numer. Anal.* **34**, 749 (2000).
- <sup>14</sup>E. Cancès and C. Le Bris, *Int. J. Quantum Chem.* **79**, 82 (2000).
- <sup>15</sup>E. Cancès, *J. Chem. Phys.* **114**, 10616 (2001).
- <sup>16</sup>P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980).
- <sup>17</sup>P. Pulay, *J. Comput. Chem.* **3**, 556 (1982).
- <sup>18</sup>K. N. Kudin, G. E. Scuseria, and E. Cancès, *J. Chem. Phys.* **116**, 8255 (2002).
- <sup>19</sup>K. N. Kudin and G. E. Scuseria, *Math. Modell. Numer. Anal.* **41**, 281 (2007).
- <sup>20</sup>L. Thøgersen, J. Olsen, D. Yeager, P. Jørgensen, P. Salek, and T. Helgaker, *J. Chem. Phys.* **121**, 16 (2004).
- <sup>21</sup>L. Thøgersen, J. Olsen, A. Kohn, P. Jørgensen, P. Salek, and T. Helgaker, *J. Chem. Phys.* **123**, 074103 (2005).
- <sup>22</sup>S. Høst, J. Olsen, B. Jansik, L. Thøgersen, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **129**, 124106 (2008).
- <sup>23</sup>S. Høst, B. Jansik, J. Olsen, P. Jørgensen, S. Reine, and T. Helgaker, *Phys. Chem. Chem. Phys.* **10**, 5344 (2008).
- <sup>24</sup>R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (Wiley, Chichester, 1980).

- ester, 1987).
- <sup>25</sup> An in-house program for QM/MM simulations (<http://www.qm4d.info>).
- <sup>26</sup> J. Nocedal, *Math. Comput.* **35**, 773 (1980).
- <sup>27</sup> D. Liu and J. Nocedal, *Math. Program. Ser. B* **45**, 503 (1989).
- <sup>28</sup> G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, *J. Chem. Phys.* **89**, 2193 (1988).
- <sup>29</sup> G. A. Petersson and M. A. Al-Laham, *J. Chem. Phys.* **94**, 6081 (1991).
- <sup>30</sup> A. D. Becke, *Phys. Rev. B* **38**, 3098 (1988).
- <sup>31</sup> A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- <sup>32</sup> C. T. Lee, W. T. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>33</sup> S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- <sup>34</sup> J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.* **102**, 939 (1980).
- <sup>35</sup> K. D. Dobbs and W. J. Hehre, *J. Comput. Chem.* **8**, 880 (1987).
- <sup>36</sup> W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- <sup>37</sup> M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees, and J. A. Pople, *J. Chem. Phys.* **77**, 3654 (1982).
- <sup>38</sup> W. R. Wadt and P. J. Hay, *J. Chem. Phys.* **82**, 284 (1985).
- <sup>39</sup> W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- <sup>40</sup> X. Zeng, H. Hu, X. Q. Hu, A. J. Cohen, and W. T. Yang, *J. Chem. Phys.* **128**, 124510 (2008).
- <sup>41</sup> A. J. Cohen, P. Mori-Sánchez, and W. T. Yang, *J. Chem. Phys.* **129**, 121104 (2008).
- <sup>42</sup> A. J. Cohen, P. Mori-Sánchez, and W. T. Yang, *Science* **321**, 792 (2008).
- <sup>43</sup> P. Mori-Sánchez, A. J. Cohen, and W. T. Yang, *Phys. Rev. Lett.* **100**, 146401 (2008).
- <sup>44</sup> C. D. Zeinalipour-Yazdi, A. L. Cooksy, and A. M. Efstathiou, *Surf. Sci.* **602**, 1858 (2008).