

Implementation of the Diagonalization-Free Algorithm in the Self-Consistent Field Procedure within the Four-Component Relativistic Scheme

Marcela Hrdá,^[a] Tomáš Kulich,^[b] Michal Repiský,^[a,c] Jozef Noga,^[a,d,e] Olga L. Malkina,^{*,[a,d]} and Vladimir G. Malkin^[a]

A recently developed Thouless-expansion-based diagonalization-free approach for improving the efficiency of self-consistent field (SCF) methods (Noga and Šimunek, *J. Chem. Theory Comput.* 2010, 6, 2706) has been adapted to the four-component relativistic scheme and implemented within the program package ReSpect. In addition to the implementation, the method has been thoroughly analyzed, particularly with respect to cases for which it is difficult or computationally expensive to find a good initial guess. Based on this analysis, several modifications of the original algorithm, refining its

stability and efficiency, are proposed. To demonstrate the robustness and efficiency of the improved algorithm, we present the results of four-component diagonalization-free SCF calculations on several heavy-metal complexes, the largest of which contains more than 80 atoms (about 6000 4-spinor basis functions). The diagonalization-free procedure is about twice as fast as the corresponding diagonalization. © 2014 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23674

Introduction

The electronic structure and molecular properties of systems containing heavy elements are influenced by relativistic effects, in most cases to an extent that prevents modeling the molecular systems accurately using perturbation theories.^[1] Therefore, the development of accurate and efficient electronic structure methods that include relativistic corrections variationally represents an important challenge of modern theoretical chemistry.^[2–7] A rigorous starting point for such a development is the four-component relativistic formalism based on the Dirac Hamiltonian.^[8]

During the last decades, much effort has been invested in the development and implementation of relativistic electronic structure models, starting from simple self-consistent field (SCF) methods, such as Dirac–Hartree–Fock^[9–15] or Dirac–Kohn–Sham (DKS),^[16–19] up to elaborate electron correlation methods (for a comprehensive review of relativistic correlation methods see Ref. [5] and references therein). Among the methods developed, DKS density functional theory represents the best compromise between cost and accuracy and it is, therefore, important to be able to perform DKS ground-state optimizations in a reliable and efficient manner. Higher computational cost and complexity of the relativistic approaches, when compared to the nonrelativistic case, is associated with the multicomponent form of the Dirac Hamiltonian whose matrix representation requires the evaluation of integrals over a complex 4-spinor basis $\{\chi\} = \{\chi^L, \chi^S\}$, where L and S denote the large and the small components, respectively. In particular, the small-component 2-spinors, $\{\chi^S\}$ lead to significant complications, as their form is governed to lowest order in c^{-2} by the restricted kinetically balanced rela-

tion^[20,21] (RKB): $\{\chi^S\} = \{(\sigma \cdot p)\chi^L\}$. In the presence of a magnetic field, an even more complicated restricted magnetically balanced condition^[22,23]: $\{\chi^S\} = \{(\sigma \cdot \pi)\chi^L\}$ must be applied to represent the coupling between basis set components to the same order (c^{-2}). These inherent complications make the efficient implementation of four-component electronic structure methods, and thus their application beyond small chemical systems, difficult.

In standard nonrelativistic and relativistic SCF methods, the stationary (Dirac–)Hartree–Fock or (Dirac–)Kohn–Sham

[a] M. Hrdá, M. Repiský, J. Noga, O. L. Malkina, V. G. Malkin
Department of Theoretical Chemistry, Institute of Inorganic Chemistry,
Slovak Academy of Sciences, SK-84536 Bratislava, Slovakia. E-mail:
olga.malkin@savba.sk

[b] T. Kulich
Department of Computer Science, Faculty of Mathematics, Physics and
Informatics, Comenius University, SK-84248 Bratislava, Slovakia

[c] M. Repiský
Department of Chemistry, The Centre for Theoretical and Computational
Chemistry (CTCC), UiT, The Arctic University of Norway, N-9037 Tromsø,
Norway

[d] J. Noga, O. L. Malkina
Department of Inorganic Chemistry, Faculty of Natural Sciences, Comenius
University, SK-84215 Bratislava, Slovakia

[e] J. Noga
Computing Centre, Slovak Academy of Sciences, SK-84545 Bratislava,
Slovakia
Contract/grant sponsor: Grant Agency of the Ministry of Education of the
Slovak Republic and Slovak Academy of Sciences VEGA; Contract/grant
number: VEGA 2/0148/13; Contract/grant sponsor: Slovak Research and
Development Agency; Contract/grant numbers: APVV-0510-12 and
APVV-0483-10; Contract/grant sponsor: Research Council of Norway (to
M.R.); Contract/grant number: 214095-F20; Contract/grant sponsor:
Norwegian Supercomputing Program; Contract/grant number: nn4654K

© 2014 Wiley Periodicals, Inc.

equations are solved iteratively. In conventional algorithms, every direct SCF iteration consists of two steps: the construction of an effective one-electron Hamiltonian (Fock matrix), and solving the related one-particle characteristic problem by diagonalization of the Fock matrix. Much effort has been spent on developing and implementing techniques reducing the computational effort and addressing the scaling issues in the SCF steps. For a comprehensive review of the acceleration techniques used in the realm of nonrelativistic theory, the reader is referred to Ref. [24].

In relativistic calculations, the most time-consuming part is typically associated with the evaluation of Coulomb (**J**) and exchange (**K**) contributions to the Fock matrix. Similarly to the nonrelativistic case, their computational complexity grows formally as $\mathcal{O}(N^4)$ due to the need to evaluate four-center electron repulsion integrals over N single-particle basis functions. Gains from the accelerating techniques are most pronounced for the Coulomb term, where the use of a density-weighted integral prescreening,^[13] similar to the one originally proposed for the nonrelativistic formalism,^[25] reduces its scaling to $\mathcal{O}(N^2)$. A very large prefactor, typically associated with the evaluation of surviving matrix elements, can be further reduced by additional techniques, such as the one proposed by Visscher, in which the Coulomb interaction between two molecular small-component charge distributions is obtained to a good approximation from tabulated atomic small-component densities.^[26] Other acceleration schemes use either pseudospectral numerical integration^[27] or the expansion of the relativistic electron density into preoptimized auxiliary basis sets, known as the density fitting approach.^[16,28–31] Although the aforementioned approaches can conveniently reduce the scaling of the Coulomb term, the evaluation of the exchange contribution using, for instance, the electron density fitting approach does not lead to the same efficiency; partly due to the importance of involving larger auxiliary basis sets to represent two-particle overlap distributions accurately.^[32] As shown by Kelley and Shiozaki on Dirac–Hartree–Fock calculations, while the use of standard nonrelativistic auxiliary basis sets introduces errors similar to the nonrelativistic case, the performance of such basis sets in relativistic calculations involving the Gaunt or the full Breit interaction gives poor results.^[31]

All these methodological advances represent an important leap forward, in particular when the acceleration techniques are combined with the recent development of modern relativistic integral programs allowing for the simultaneous evaluation of electron repulsion integrals over the large- and small-component basis (see work of Yanai et al.,^[33] Grant and Quiney,^[34] Repisky,^[30,35] or Shiozaki^[36]). As shown very recently by several groups, all the improvements reduce the computational cost of four-component SCF methods to a level where these calculations can be performed routinely on systems containing more than hundred atoms and/or a few thousands 4-spinor basis functions.^[31,37,38] In this case, the computational bottleneck is progressively shifted toward the matrix diagonalization. Moreover, as diagonalization is hard to parallelize well, it becomes a serious bottleneck in large-scale quantum chem-

istry simulations that are run on many processors. It is also worth noting that diagonalization becomes the bottleneck of the entire fully relativistic calculations much earlier compared to the nonrelativistic formalism, as the relativistic Fock matrix is complex and its dimension is four times larger. It is, therefore, desirable to implement efficient four-component SCF ground-state optimization without invoking the diagonalization step.

Recently, Noga and Šimunek^[39] proposed a promising diagonalization-free algorithm that can be used in SCF procedures. In contrast to conventional diagonalization, the pertinent algorithm is easily parallelizable and it provides much faster calculation as well. Moreover, it was reported that this approach can reduce the number of iterations needed to reach the SCF convergence,^[40] which would improve the efficiency of SCF approach in general, that is, even in cases when the time spent on the diagonalization step is negligible.

The main purpose of this work has been to formulate and implement the pertinent diagonalization-free approach in the fully relativistic four-component scheme. A modified version of this approach was successfully implemented at the one-component^[39] and two-component wave function levels, already.^[40] Nevertheless, it was noted^[39,40] that the algorithm may be unstable when one starts the calculation from a poor initial guess. Obtaining a good initial guess in relativistic calculations is often difficult. Moreover, the SCF convergence becomes more problematic due to the common presence of many near-degenerate orbitals in heavy-element compounds (the main subject of relativistic calculations) and additional variational freedom. Thus, the stability of the algorithm in the relativistic framework is critical. Beside our main goal, in this article, we provide somewhat deeper analysis of the stability and numerical efficiency of the diagonalization-free algorithm. Based on this analysis, we propose further modifications of the approach. Different benchmark calculations performed at the one-component level allowed us to tune the efficiency of the diagonalization-free code toward still better performance. Finally, we present our implementation of the improved algorithm at the four-component level in the ReSpec program.^[41]

The article is organized as follows. In Methodology Section, we briefly review the Thouless expansion-based diagonalization-free approach, propose further improvements of the original approach toward better efficiency and stability and generalize it to the four-component formalism. Numerical results are presented in the next section (Benchmark Four-Component Calculations). Finally, concluding remarks are drawn.

Methodology

Solution of the Hartree–Fock and Kohn–Sham equations is most straightforwardly obtained via diagonalization of the Fock matrix, which is generally expressed in terms of a nonorthogonal one-particle basis set χ of a finite dimension n . Diagonalization provides n orthogonal one-particle eigenstates (molecular orbitals—MOs), and, for an N -electron system, the

N lowest-energy eigenvectors are used to construct the N -electron ground-state Slater determinant. Consequently, the one-particle density matrix is constructed using those eigenvectors in each SCF cycle, whereas the new density matrix is used to construct the new Fock matrix. This procedure is repeated until optimal MOs are obtained.

In fact, what we need to have updated in each iteration is the mere density matrix that determines the final energy. The diagonalization of the Fock matrix is not the only way how this update can be done. Several diagonalization-free solutions have been suggested. For the state of the art in this field, we refer the reader to recent articles.^[42–44] In this work, we have closely followed the procedure suggested by one of us,^[39] due to its conceptual simplicity and an attractive feature of the approach that makes it different from most of the other diagonalization-free methods, namely that the calculated density matrix is exactly idempotent in any step and it is naturally N -representable.

To introduce modifications within the four-component scheme, we have first analyzed in more detail the one-component version and also performed some related numerical experiments targeted toward stability and computational efficiency improvements, which are summarized after a brief recapitulation of the core of the pertinent theory. In the last subsection of this section, we shall introduce the necessary modifications of the algorithm to adapt it to the relativistic four-component scheme.

Using the Thouless theorem

Let us start by fixing a reference MO basis $\varphi_p, p=1, \dots, n$, where n is the number of MOs. This MO basis is orthonormal and most frequently the individual MOs are given as linear combinations of computational basis functions from a nonorthogonal set χ , using a transformation matrix \mathbf{C}^{ref} , that is,

$$\varphi = \mathbf{C}^{\text{ref}} \chi. \quad (1)$$

The reference N -electron Slater determinant $|\Phi\rangle$ is composed of the N lowest-energy reference MOs $\varphi_i, i=1, \dots, N$:

$$|\Phi\rangle = |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle. \quad (2)$$

We will call these orbitals occupied (o) and we will use letters i, j, \dots to denote them. The complementary orbitals will be called virtual (v) and denoted by a, b, \dots . Letters p, q, \dots will refer to arbitrary reference orbitals.

Let us define an operator $a_i^a = a_a^\dagger a_i$, where a_i is the annihilation operator of state $|\varphi_i\rangle$ and a_a^\dagger is the creation operator of the state $|\varphi_a\rangle$. Hence,

$$a_i^a |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle = |\varphi_1 \dots \varphi_a \dots \varphi_N\rangle. \quad (3)$$

The global single excitation operator T_1 is a linear combination of operators a_i^a for all possible values of i and a :

$$T_1 = \sum_{i,a} t_a^i a_i^a, \quad (4)$$

where t_a^i are the amplitudes related to the excitation from the state φ_i to the state φ_a .

According to the Thouless theorem,^[45] any N -particle Slater determinant, which is not orthogonal to $|\Phi\rangle$, can be written in an exponential form

$$|\Psi\rangle = e^{T_1} |\Phi\rangle. \quad (5)$$

Conversely, the actions of an exponential operator e^{T_1} on an N -particle Slater determinant $|\Phi\rangle$ will always produce an N -particle Slater determinant. Let \mathbf{t} denote the amplitude matrix of dimensions o, v containing elements t_a^i . Hence, instead of searching for optimal MOs that can be used to construct the optimal Slater determinant, one can find optimal \mathbf{t} such that the Slater determinant (5) is the optimal one while the reference MOs are not changed. The key point of the new diagonalization-free approach is the fact that with the given \mathbf{t} , the one-particle density matrix can be exactly evaluated,^[46] despite the fact that the density matrix expansion constitutes an infinite series. In the Hartree–Fock case, solution for \mathbf{t} follows from the minimization of the expectation value energy functional using wave function (5) w.r.t. the amplitudes of T_1^\dagger , which gives rise to an analogy of the Brillouin condition.^[46] Equivalently, the solution for \mathbf{t} is obtained by step by step update of \mathbf{t} that follows from the minimization of the expectation value of the Fock (or, in case of DFT, Kohn–Sham) operator.^[39] Such a procedure should converge when $|\Phi\rangle$ is a dominant determinant in the expansion of eq. (5), that is, when the reference basis is not too far from the optimal Hartree–Fock orbitals. In case of the Kohn–Sham approach, a similar statement can be formulated considering a reference system of noninteracting particles.

Construction of the one-particle density matrix

If the amplitudes t_a^i are all equal to zero, the density matrix \mathbf{D}^{ref} written in the reference MO basis can be expressed as

$$\mathbf{D}^{\text{ref}} = \begin{pmatrix} \mathbf{1}_{oo} & \mathbf{0}_{ov} \\ \mathbf{0}_{vo} & \mathbf{0}_{vv} \end{pmatrix}, \quad (6)$$

where the subscripts oo, ov, vo, and vv denote the pertinent occupied-occupied, occupied-virtual, virtual-occupied, and virtual-virtual blocks of the full density matrix. Hence, $\mathbf{D}_{oo}^{\text{ref}} = \mathbf{1}_{oo}$ is a unit matrix, whereas all other blocks of \mathbf{D}^{ref} are zero matrices.

The density matrix \mathbf{D} corresponding to a general non-zero amplitude matrix \mathbf{t} can be written as a sum of \mathbf{D}^{ref} and a correction matrix \mathbf{X} ,

$$\mathbf{D} = \begin{pmatrix} \mathbf{1}_{oo} + \mathbf{X}_{oo} & \mathbf{X}_{ov} \\ \mathbf{X}_{vo} & \mathbf{X}_{vv} \end{pmatrix}. \quad (7)$$

Even though \mathbf{X} is given as an infinite expansion following from the exponential structure of the wave function eq. (5), it turns out that its explicit evaluation is rather straightforward using the diagrammatic technique.^[46] When one first constructs the matrices

$$\mathbf{x}_{oo} = -\mathbf{t}\mathbf{t}^\dagger, \quad (8)$$

$$\mathbf{x}_{vv} = -\mathbf{t}^T \mathbf{t}, \quad (9)$$

the individual blocks of the correction matrix \mathbf{X} read:

$$\mathbf{x}_{oo} = \sum_{m=1}^{\infty} \mathbf{x}_{oo}^m, \quad (10)$$

$$\mathbf{x}_{vv} = -\sum_{m=1}^{\infty} \mathbf{x}_{vv}^m, \quad (11)$$

$$\mathbf{x}_{ov} = \mathbf{x}_{vo}^T = \mathbf{t} + \mathbf{x}_{oo} \mathbf{t}. \quad (12)$$

Because \mathbf{x}_{vv} at the same time is given as

$$\mathbf{x}_{vv} = -\mathbf{t}^T \mathbf{x}_{ov} \quad (13)$$

the whole \mathbf{X} matrix is easily determined as soon as we have its oo-block. In fact, it is convenient to calculate the oo-block of the density matrix instead,^[39] either using a recurrent microiteration scheme:

$$\mathbf{D}_{oo}^{(0)} = \mathbf{1}_{oo} \quad (14)$$

$$\mathbf{D}_{oo}^{(l+1)} = \mathbf{1}_{oo} + \mathbf{D}_{oo}^{(l)} \mathbf{x}_{oo}, \quad (15)$$

or directly by matrix inversion

$$\mathbf{D}_{oo} = (\mathbf{1}_{oo} - \mathbf{x}_{oo})^{-1}. \quad (16)$$

In Ref. [39], the recurrent scheme was preferred because it merely involved simple matrix multiplications. Nevertheless, it was suggested that the approach with matrix inversion can be useful when the recurrence relation eq. (15) converges slowly.

Calculation of the remaining blocks of the density matrix \mathbf{D} follows easily from eqs. (10) to (13):

$$\mathbf{D}_{ov} = \mathbf{x}_{ov} = \mathbf{D}_{oo} \mathbf{t}, \quad (17)$$

$$\mathbf{D}_{vo} = \mathbf{x}_{vo} = \mathbf{t}^T \mathbf{D}_{oo}, \quad (18)$$

$$\mathbf{D}_{vv} = \mathbf{x}_{vv} = \mathbf{t}^T \mathbf{D}_{oo} \mathbf{t}. \quad (19)$$

Updating the t-amplitudes

The t-amplitudes have to minimize the expectation value of the Fock operator for the state given by eq. (5)

$$\frac{\partial \text{Tr}[\mathbf{F}\mathbf{D}(\mathbf{t}^T, \mathbf{t})]}{\partial \mathbf{t}} = 0, \quad (20)$$

where \mathbf{D} is the density matrix of the state eq. (5) constructed from the t-amplitudes as described above. At this step in the given iteration, the Fock matrix \mathbf{F} can be treated as fixed and independent of \mathbf{t} . It can be shown^[39] that eq. (20) results in a matrix equation:

$$(\mathbf{1}_{oo} + \mathbf{x}_{oo})[\mathbf{F}_{ov}(\mathbf{1}_{vv} - \mathbf{x}_{vv}) - \mathbf{F}_{oo} \mathbf{x}_{ov}] + \mathbf{x}_{ov}[\mathbf{F}_{vv}(\mathbf{1}_{vv} - \mathbf{x}_{vv}) - \mathbf{F}_{vo} \mathbf{x}_{ov}] = \mathbf{0}_{ov} = \tilde{\mathbf{F}}_{ov}. \quad (21)$$

It is not easy to obtain the whole matrix \mathbf{X} directly from eq. (21). A simple and algorithmically favorable approach is to

update \mathbf{t} and subsequently evaluate the blocks of \mathbf{D} as described above. Realizing that the leading term in the expansion of \mathbf{x}_{ov} is \mathbf{t} and using some other approximations (see Appendix B) a simple update of the t-amplitudes in the $(k+1)$ -st SCF iteration has been suggested^[39]:

$$t_a^{i(k+1)} = t_a^{i(k)} + \frac{\tilde{F}_a^{i(k)}}{f_i^i - f_a^a} \quad (22)$$

starting with

$$t_a^{i(0)} = 0. \quad (23)$$

In (22), f_i^i, f_a^a are diagonal matrix elements of the original Fock matrix \mathbf{f} in the reference basis (i.e., $\mathbf{F}^{(0)}$) and $\tilde{\mathbf{F}}^{(k)}$ is defined as

$$\tilde{\mathbf{F}}^{(k)} = \mathbf{Q}^{(k)T} \mathbf{F}^{(k)} \mathbf{Q}^{(k)}, \quad (24)$$

where

$$\mathbf{Q}^{(k)} = \mathbf{1} + \begin{pmatrix} \mathbf{x}_{oo}^{(k)} & -\mathbf{x}_{ov}^{(k)} \\ \mathbf{x}_{vo}^{(k)} & -\mathbf{x}_{vv}^{(k)} \end{pmatrix}. \quad (25)$$

Note that the ov-block of $\tilde{\mathbf{F}}^{(k)}$ is in fact the residual of eq. (21) in the k th iteration and approaches zero at the convergence limit.

One merely needs the density matrix in the SCF calculations, irrespective how it was obtained. Nonetheless, if necessary one can construct "optimal" orthonormalized MOs from \mathbf{C}^{ref} using the transformation matrix \mathbf{Q} as^[39]

$$\mathbf{C}_o = \mathbf{C}^{\text{ref}} \mathbf{Q}_o \mathbf{Q}_{oo}^{-1/2}, \quad (26)$$

$$\mathbf{C}_v = \mathbf{C}^{\text{ref}} \mathbf{Q}_v \mathbf{Q}_{vv}^{-1/2}, \quad (27)$$

where

$$\mathbf{Q}_o = \begin{pmatrix} \mathbf{Q}_{oo} \\ \mathbf{Q}_{vo} \end{pmatrix}, \mathbf{Q}_v = \begin{pmatrix} \mathbf{Q}_{ov} \\ \mathbf{Q}_{vv} \end{pmatrix}. \quad (28)$$

Improving the efficiency and the stability

Baldes et al.^[40] found that the stability of the original algorithm^[39] strongly depended on the starting reference MOs. If the reference MOs were too distant from the solution of the SCF equation, very slow or no convergence of the SCF iterations and microiterations was observed. To solve this problem, a hybrid scheme was suggested and implemented^[40] using the following idea: monitor the largest element $\max(\mathbf{t})$ and apply the diagonalization-free step only if the pertinent amplitude does not exceed a threshold value set to 0.1. Otherwise, perform one diagonalization and obtain new and supposedly better reference MOs. The related threshold was empirically determined from a large set of molecules, as a pragmatic value. Baldes et al.^[40] also suggested using a slightly different update of the t-amplitudes

$$t_a^{i(k+1)} = t_a^{i(k)} + \frac{\tilde{F}_a^{i(k)}}{\tilde{F}_i^{i(k)} - (\tilde{F}_a^{a(k)} + \Delta)}, \quad (29)$$

This equation differs from the original eq. (22) in two ways: (i) the diagonal elements of the Fock matrix from the current iteration are used in the denominator; (ii) $\Delta > 0$ is used to shift the virtual orbitals to higher energies, a practice common also in conventional SCF algorithms. One has to note, however, that while for an update via eq. (22) one merely needs a partial transformation of the Fock matrix to its **ov** block, in eq. (29), the transformation must be performed also for the diagonal elements.

Calculations on a test set of about 400 mostly inorganic molecules have shown^[40] that the hybrid algorithm with modified eq. (29) was stable and reliable, while somewhat faster convergence was achieved.

Our experience using a straightforward pilot implementation of the diagonalization-free algorithm in the ReSpect code corresponded to that reported in the aforementioned work. There was essentially no convergence problem when the t-amplitudes were small, but starting with “bad” reference MOs, too many (e.g., even hundreds of) SCF iterations were needed to reach convergence, or there were no signs of SCF convergence at all. In addition, there were too many microiterations needed to converge eq. (15) or the microiterations even diverged. Because we intend to apply the diagonalization-free procedure to speed up the four-component SCF calculations where obtaining a good initial guess is more problematic, the stability of the algorithm is crucial. We have, therefore, reconsidered the algorithm with the goal of increasing the reliability of the algorithm when one starts from worse initial conditions. This implies that the initial amplitudes are larger, which causes a slow or no convergence of the recurrent relation eq. (15). In the initial article using this method,^[39] it was noted that the microiterations converge if the matrix \mathbf{x}_{oo} defined by eq. (8) has all the eigenvalues smaller than 1, because the series in eq. (10) is then converging. It turned out that the density matrix can be evaluated even if this condition is not fulfilled (see Appendix A).

Therefore, substituting the microiterations with matrix inversion completely eliminates the problem with a possible instability in the course of the density matrix calculation. In the

Table 1. Relative timings for different matrix operations with respect to the matrix multiplication of equal dimensionality that is taken as a reference.

Operation	MKL routine	Relative timing CPU/CPU(dgemm)
Multiplication	dgemm	1.00
Diagonalization	dsyev	10.88
PLU decomposition	dgetrf	0.35
Inversion	dgetrf and dgetri	1.04

The subroutine `dsyev` was used to compute both eigenvalues and eigenvectors. In all tests, we used routines of Intel® Math Kernel Library. See Supporting Information for further details.

original work,^[39] it was assumed that the computational efficiency of the matrix inversion was inferior with respect to the matrix multiplications needed in eq. (15). Our current numerical experiments suggest that this assumption was not correct, as illustrated in Table 1. There, we compare the CPU timings for basic matrix operations involved in different steps of the algorithm, always reported relative to the matrix multiplication. The formal computational complexity for all of the listed operations scales as $\mathcal{O}(n^3)$ with the matrix dimension n . As documented in Supporting Information, these relative timings are indeed independent of the matrix dimensions for a wide range of dimensions that are of interest to us. A single matrix inversion approximately corresponds to one matrix multiplication of the same sized matrices. Taking into account that in the microiterative procedure eq. (15), we usually need several consecutive matrix multiplications, matrix inversion should be preferable.

Even though the density matrix can be accurately calculated with any set of t-amplitudes, it is unlikely that the whole SCF procedure will converge with t-amplitudes larger than unity, that is, when the reference determinant is not the dominant one. Nevertheless, we have encountered such cases during our numerical experiments. In those cases, however, we had to use the matrix inversion [eq. (16)] to calculate the oo-block of the density matrix. For illustration, such examples are given in Table 2. Nevertheless, even though the solution could be accomplished, the table shows that the convergence of the SCF procedure was too slow. Hence, allowing diagonalizations to improve the reference basis would be more efficient.

Table 2. Examples of molecular calculations where the recurrent formula in eq. (15) failed to converge, but using eq. (16) resulted in a correct solution.

System	Basis	Number of diag. iterations	max (t) (all iterations)	max (t) (last iteration)	Number of SCF iterations
H ₂ O	6-31G ^[47]	1	1.103	1.103	254
HF	STO-3G ^[48]	1	0.683	0.683	247
N ₂ O	STO-3G ^[48]	5	1.622	0.912	34
CBr ₄	STO-3G ^[48]	4	1.129	1.129	103
CH ₄	STO-3G ^[48]	1	1.128	1.127	192
CO	upc-0 ^[49]	2	2.382	1.438	209
HBr	STO-3G ^[48]	3	1.430	1.430	490

Reference MOs were obtained by starting with bare nucleus initial guess and performing several diagonalization steps at the beginning of the SCF calculation (see column number of diag. iterations). After that point, the pure diagonalization-free approach was applied. We present the number of all SCF iterations needed to converge to the correct MOs, the maximum t-amplitude in the last SCF iteration and the maximum of max (t) over all SCF iterations.

Table 1 also demonstrates that irrespective of its poor potential for parallelization, performing the diagonalization is much less effective than carrying out the necessary matrix multiplications in the Thouless expansion approach. As for the computational complexity, we note that unlike the diagonalization that has to be performed with the matrix of dimension n corresponding to the number of basis functions (i.e., scaling as $\mathcal{O}(n^3)$), matrix multiplications in the microiteration procedure and inversion of the matrix are related to much smaller matrices whose dimensions are equal to the number of occupied orbitals N . Moreover, transformation of the Fock matrix [eq. (24)] is only needed for the **ov** block, that is, its computational complexity scales as $\mathcal{O}(Nn^2)$. Therefore, we have investigated alternative criteria that could still give rise to decreasing the number of diagonalization steps within the hybrid approach.

Both eqs. (22) and (29) are approximate solutions of eq. (21). A motivation for just using these simple relations that would correspond to the first iteration in the iterative solution of the nonlinear eq. (21) is twofold: (i) the update is extremely simple from the computational point of view; (ii) the transformation matrix **Q** in eq. (24) is fully compatible with the density matrix that was used to create the Fock matrix, that is, both are stemming from the same **t**. It can hence be assumed that such an approach is well balanced.

However, as shown in Appendix B, in case of a poor initial guess for the MOs, the approximations made in eq. (22) or (29) may be too rough, causing slow convergence or instability of the diagonalization-free SCF procedure. These problems can be resolved by applying the diagonalization-free step within the pragmatic hybrid approach.^[40] Unfortunately, as follows from Appendix B, it is difficult to predict whether the given MOs are good or not. It seems there is a general trend: the diagonalization-free approach works well with small **t**-amplitudes and it often fails for higher ones. One can choose tight and rigid criteria for **t**-amplitudes, making the hybrid approach stable for a large number of systems, but such criteria might sometimes be too restrictive, causing unnecessary increase of the diagonalization steps and thus deterioration of the computational efficiency. Hence, a more flexible practical approach is desired.

The total error caused using approximate **t**-updating equations can be easily estimated numerically. As eq. (21) is precise, one can check whether the new (updated) **t**-amplitudes satisfy this equation reasonably well in comparison to the old **t**-amplitudes.

Equation (21) requires that $\tilde{\mathbf{F}}_{\text{ov}} = \mathbf{Q}_0^\dagger \mathbf{F} \mathbf{Q}_v$ is equal to 0. The ratio of the euclidean norms of this matrix block before and after updating **t** in the k th iteration, namely

$$r = \|\mathbf{Q}_0^{(k+1)\dagger} \mathbf{F}^{(k)} \mathbf{Q}_v^{(k+1)}\| / \|\mathbf{Q}_0^{(k)\dagger} \mathbf{F}^{(k)} \mathbf{Q}_v^{(k)}\|. \quad (30)$$

can be, therefore, viewed as a measure of the inaccuracy of the update of **t** and can be used to indicate the necessity for applying the diagonalization step. Smaller ratios justify the update of **t** whereas large values indicate that a diagonaliza-

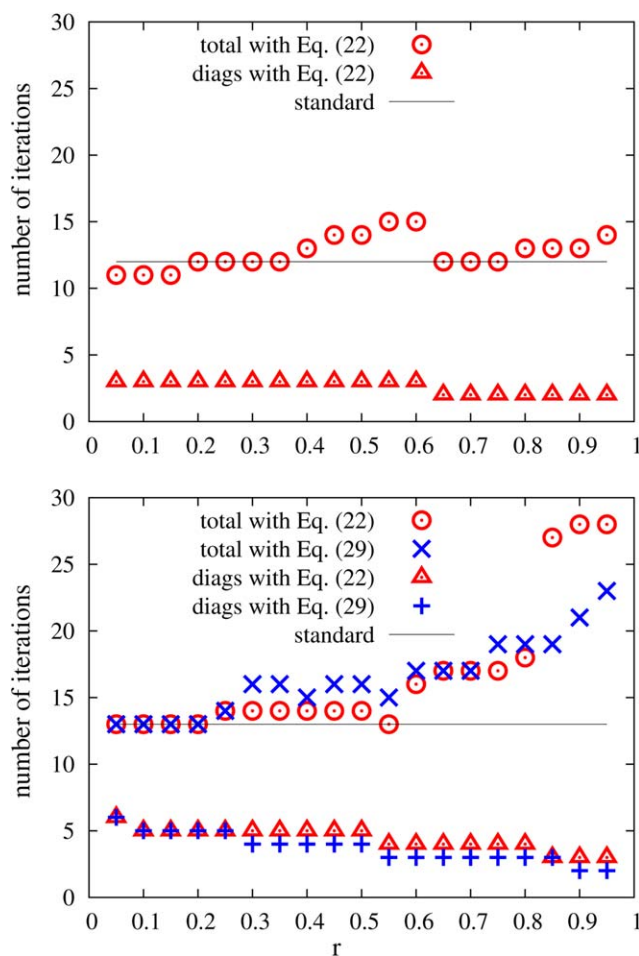


Figure 1. Typical numbers of SCF iterations required with varying threshold for r and using different update for the **t**-amplitudes; iterations with diagonalizations within the hybrid approach are distinguished; “standard” denotes the conventional diagonalization procedure for comparison. Bottom panel: C₆H₆, basis upc-1, bare nucleus initial guess; top panel: P₄O₁₀, basis upc-1, EHT initial guess.

tion step should be performed. To set a proper threshold for this quantity, numerical experience was needed.

To test and tune the procedure, we have performed calculations on a variety of molecules using both eqs. (22) and (29) to update the **t**-amplitudes and varying the threshold for both r and the **t**-amplitudes. As follows from Supporting Information data, in most cases there is not a significant difference in the convergence behavior either using the aforementioned criterion or the **t**-amplitude value threshold. However, when the threshold for the **t**-amplitude criterion is set to higher values, the probability of falling into the unstable regime becomes higher. In contrast, the criterion [eq. (30)] gives rise to a stable convergence within a wide range of ratios, spanning from about 0.05 to 0.95. Note that values over unity may eventually lead to divergence. A typical convergence pattern for a varying threshold within the latter range is displayed in Figure 1, where we compare the total number of iterations within the hybrid approach with the number of iterations in a “standard” procedure based on the diagonalization of the Fock matrix. At the same time, numbers of prompted diagonalizations within the hybrid approach are given. In the bottom panel, we

present results for the benzene molecule using a bare nucleus initial guess, representing a rather poor start for our procedure. Here, we show the convergence pattern with employing either eq. (22) or (29) for an update of the t -amplitudes, while no denominator damping was applied. Despite the second approach appearing preferable from the theoretical point of view, no significant difference in the convergence pattern has been observed between the two approaches. This statement turned out to be statistically valid in all our computations, hence eq. (22) is taken as the default in our code due to its simpler form. Still, the denominator damping can be switched on in the case of poor convergence.

In the top panel of Figure 1, the convergence behavior is displayed for a more complex example of P_4O_{10} (optimized geometry by Demissie and Repisky, unpublished) now from the EHT initial guess and only using the default update of the t -amplitudes.

These two examples represent typical convergence behaviors for a broader set of molecules that we have used for testing. The number of prompted diagonalizations is quite insensitive to the threshold value within the investigated range; yet this number is decreasing slightly with higher values of the threshold. Conversely, the total number of iterations increases at higher values of the threshold. Hence, a threshold value around 0.5 seems to be optimal.

Generalization to four-component wave function

The relativistic four-component SCF differs from the one-component SCF in two aspects with respect to the aforementioned diagonalization-free method, namely:

1. Four-component Dirac–Hartree–Fock and DKS equations are complex, hence one has to deal with complex arithmetic. Operators and the density matrix are still represented by Hermitian matrices but unlike in one-component scheme, those are no longer real and symmetric. “Bra” and “ket” vectors are not just mutually transposed, but indeed Hermitian conjugates. The same applies to transformation matrices involved in unitary transformations.
2. In four-component calculations, one has to deal with separate basis sets for the small and the large components. We will denote the size of these bases as s and l , respectively. In four-component calculation, s lowest-energy eigenvalue-eigenvector pairs do not correspond to real electron states. Therefore, they are not occupied. As a consequence, the oo -block of density matrix written in MO basis (ordered with respect to increasing one-electron energies) is not located in its top-left corner, but somewhere on the diagonal of the density matrix:

$$\begin{pmatrix} \mathbf{0}_{ss} & & \mathbf{0}_{sl} \\ \mathbf{0}_{ls} & \mathbf{1}_{NN} & \mathbf{0}_{N(I-N)} \\ & \mathbf{0}_{(I-N)N} & \mathbf{0}_{(I-N)(I-N)} \end{pmatrix} \quad (31)$$

For the special case of using restricted kinetic balance (RKB) basis sets^[20,21] for the small component, as is the case of the four-component ReSpect code, the sizes of small and large

component bases are equal: $s=l=2n$, where n is the size of the pertinent one-component basis set.

Fortunately, both aforementioned issues can be easily resolved. Suppose, we have a fixed reference MO basis $\mathcal{B} = \varphi_1, \dots, \varphi_{s+l}$ ordered in accordance with increasing energy obtained in an analogous way as in the one-component calculation. Here, we summarize our algorithm as it is implemented in the current version of ReSpect. We are aware that some steps could be formally simplified, but at the moment we tried to intervene into the existing code as little as possible. A diagonalization-free SCF iteration consists of the following steps:

1. Obtain Fock matrix from the density matrix in the usual way. Transform the Fock matrix to the reference MO basis \mathcal{B} .
2. Transform Fock matrix \mathbf{F} from basis \mathcal{B} to basis $\mathcal{B}' = \varphi_{s+1}, \varphi_{s+2}, \dots, \varphi_{s+l}, \varphi_1, \dots, \varphi_l$. This can be efficiently performed as reordering of matrix blocks. For example, transformation of matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{pmatrix} \quad (32)$$

from basis \mathcal{B} to basis \mathcal{B}' results in

$$\mathbf{M}' = \begin{pmatrix} \mathbf{M}_4 & \mathbf{M}_3 \\ \mathbf{M}_2 & \mathbf{M}_1 \end{pmatrix}. \quad (33)$$

3. Use equations analogous to the one-component case to update the t -amplitudes via eq. (22) or (29) and calculate a new density matrix. The only difference in the equations is that all matrix and vector transpositions are replaced by Hermitian conjugations, hence eq. (24) now reads:

$$\tilde{\mathbf{F}} = \mathbf{Q}^\dagger \mathbf{F} \mathbf{Q}, \quad (34)$$

and

$$\mathbf{Q} = \begin{pmatrix} \mathbf{D}'_{oo} & -\mathbf{D}'_{ov} \\ \mathbf{D}'_{vo} & \mathbf{I}_{vv} - \mathbf{D}'_{vv} \end{pmatrix}. \quad (35)$$

\mathbf{D}'_{oo} , \mathbf{D}'_{ov} , \mathbf{D}'_{vo} , and \mathbf{D}'_{vv} are blocks of density matrix in basis \mathcal{B}' from the previous iteration. Calculate the new density matrix

$$\mathbf{D}' = \begin{pmatrix} \mathbf{D}'_{oo} & \mathbf{D}'_{ov} \\ \mathbf{D}'_{vo} & \mathbf{D}'_{vv} \end{pmatrix} \quad (36)$$

as follows:

$$\mathbf{D}'_{oo} = (\mathbf{I}_{oo} + \mathbf{t} \mathbf{t}^\dagger)^{-1} \quad (37)$$

$$\mathbf{D}'_{ov} = \mathbf{D}'_{oo} \mathbf{t} \quad (38)$$

$$\mathbf{D}'_{vo} = \mathbf{t}^\dagger \mathbf{D}'_{oo} \quad (39)$$

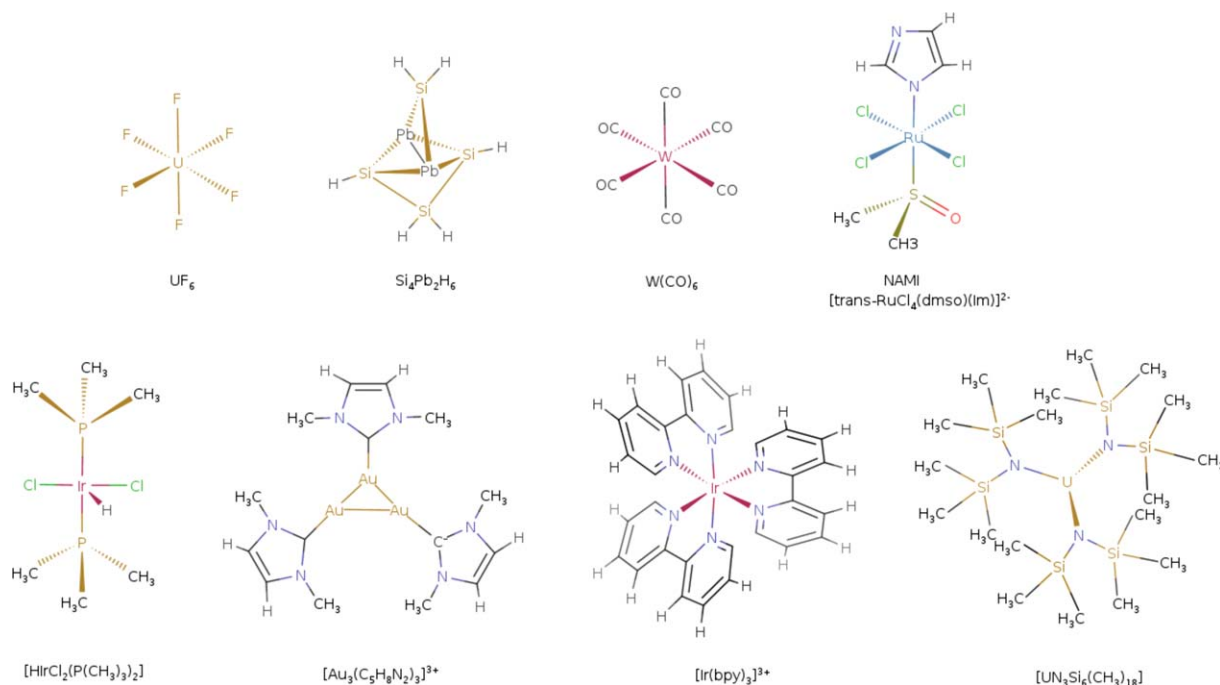


Figure 2. Sketches of the molecular structures used in testing the four-component diagonalization free algorithm. Detailed geometries are given in Ref. [38].

$$\mathbf{D}'_{vv} = \mathbf{t}^\dagger \mathbf{D}'_{oo} \mathbf{t} \quad (40)$$

4. Transform density matrix \mathbf{D}' from \mathcal{B}' to \mathcal{B} (using matrix reordering again).

The major part of these equations can be obtained by a straightforward generalization of the equations for the one-component algorithm. The only issue that needs justification is the reordering of the reference MO basis. We can do this, as in the Thouless expansion approach one never uses the fact that the reference MO basis is ordered in some specific way.

Computationally, reordering of the matrix blocks scales as $\mathcal{O}(n^2)$. This makes the second and fourth steps negligible in comparison to the third step, which implies that the considerations related to the computational efficiency for the one-component scheme are fully applicable to the four component algorithm.

Benchmark Four-Component Calculations

To illustrate the performance of the diagonalization-free approach, we carried out four-component KS-DFT ground-state optimizations with and without the diagonalization-free algorithm, and the approach was tested on a benchmark set that consisted of eight complexes containing heavy metal centers. The molecular systems were chosen to be of different size, chemical importance, and bonding situations. Figure 2 shows sketches of the calculated molecules whose detailed geometries are published elsewhere.^[38] All benchmark calculations have been carried out with the BLYP exchange-correlation functional,^[50,51] which was evaluated numerically on an adaptive molecular grid (program default). As basis sets, we used the uncontracted all-electron Dyall's basis of VDZ

quality for heavy centers (Au, Ir, Pb, Ru, U, and W),^[52] and the uncontracted cc-pVDZ basis for the light elements (C, Cl, F, H, N, O, P, S, and Si).^[53] In all calculations, we used the norm of the DIIS vector smaller than 10^{-6} as the SCF convergence criteria. Initial guess was obtained from nonrelativistic calculations. The reported calculations were performed with a local version (3.3.0) of ReSpect program.^[41]

Results are collected in Table 3. We note that in the majority of cases the total number of SCF iterations when using the

Table 3. Comparison of the number of SCF iterations required to reach convergence at the four-component (mDKS BLYP) level using standard diagonalization (third column) and the hybrid approach (fourth column).

System	nbf	Number of iterations		
		Standard	Hybrid	Speedup
UF ₆	445	13	15/1	2.1
W(CO) ₆	516	12	12/1	2.1
NAMI	610	12	12/1	2.2
Si ₄ Pb ₂ H ₆	626	15	15/1	2.1
HfCl ₂ (PMe ₃) ₂	657	11	11/1	2.1
[Ir(bpy) ₃] ³⁺	1308	12	12/1	2.2
[Au ₃ (C ₅ H ₈ N ₂) ₃] ³⁺	1326	24	42/2	2.1
UN ₃ Si ₆ Me ₁₈	1459	21	23/1	2.2

For the hybrid approach, two numbers are given: the total number of iterations and the number of iterations with standard diagonalization. Second column (nbf) describes the basis set size for one component, that is, the size of four-component complex Fock matrix was $(4 \times \text{nbf})^2$. The last column shows the ratio between the time required for one standard diagonalization and the time taken for running one diagonalization-free procedure. We do not include overall timings, as those also heavily depend on the calculation of the Fock matrix itself (that in turn depends on the used method and in case of DFT, on the exchange correlation functional), which is beyond the scope of this article.

Table 4. Comparison of the number of SCF iterations required to reach convergence of $[\text{Au}_3(\text{C}_5\text{H}_8\text{N}_2)_3]^{3+}$ complex at the four-component (mDKS BLYP) level using different uncontracted basis sets and either standard diagonalization (fourth column) or the hybrid approach (fifth column).

nbf	Basis sets		Number of iterations		Speedup
	Au	Light elements	Standard	Hybrid	
1326	dyall-vdz	cc-pVDZ	24	42/2	2.1
1611	dyall-aug-vdz	aug-cc-pVDZ	27	31/2	2.2
2634	dyall-aug-vtz	aug-cc-pVTZ	26	33/2	2.4

For the hybrid approach, two numbers are given: the total number of iterations and the number of iterations with standard diagonalization. The first column (nbf) describes the basis set size for one component, that is, the size of the four-component complex Fock matrix was $(4 \times \text{nbf})^2$. The second and third columns specify the basis sets used and the last column shows the ratio between the time required for one standard diagonalization and the time taken running one diagonalization-free procedure (but see note on Table 3).

standard and hybrid approaches was the same with the exception of the three complexes containing the heaviest atoms (uranium, or directly bonded three gold atoms). Probably for the latter the nonrelativistic guess was not as good as for the other complexes. Also, we have to note that we have not yet optimized the use of DIIS within the new algorithm. As demonstrated, the average speedup of the diagonalization step was more than a factor of two when replaced by a diagonalization-free alternative (including all the necessary transformations). In cases where diagonalization is the bottleneck of the calculation, this reduction is more than enough to even compensate for a slightly higher number of SCF iterations. It is worth mentioning that the speedup was achieved without any parallelization, but the present diagonalization-free approach is already suited for running on distributed memory parallel computers employing for instance the ScaLAPACK environment.^[54]

To explore the convergence characteristics of the diagonalization-free approach with increasing the basis set quality, we performed additional test calculations of the gold complex, $[\text{Au}_3(\text{C}_5\text{H}_8\text{N}_2)_3]^{3+}$. The complicated electronic structure of the system makes it an excellent candidate for such an analysis. We performed additional calculations where the original valence double-zeta quality basis was (a) augmented by diffuse functions (dyall-aug-vdz basis on the gold atoms and uncontracted aug-cc-pvdz basis on the light elements), and (b) extended to triple-zeta quality (dyall-aug-vtz basis on the gold atoms and uncontracted aug-cc-pvtz basis on the light elements). The results of these test calculations are presented in Table 4. Interestingly, in the calculations with extended bases, the total number of iterations for the diagonalization-free algorithm was reduced while it was slightly increased for the standard procedure. Checking whether this correlation between iterations and basis size holds in the general case would require extensive testing which is beyond the scope of the present work.

Conclusions

The optimized Thouless expansion method recently applied in independent particle model algorithms at both the nonrelativistic^[39] and relativistic two-component scheme^[40] has been extended to four-component SCF methods and imple-

mented in the program suite ReSpect. As in the two-component case, we have implemented a hybrid approach switching on a diagonalization step as soon as the monitoring of selected values indicates a potentially slow convergence or even divergence of the SCF calculation. Such a step is assumed to improve the reference determinant for the subsequent new Thouless expansion, that is, the convergence toward the optimal Thouless expansion should be significantly improved as well.

We have suggested improved calculations of the density matrix, when a single step calculation of the occupied-occupied block via the inversion [eq. (16)] proved to be computationally more efficient than its calculation via a recurrent (microiterative) procedure [eq. (15)]. Moreover, we have proved that the density matrix can be effectively evaluated for arbitrary values of the excitation amplitudes. Indeed, we have also shown examples when optimal solution for the Thouless expansion could be achieved even if the excitation amplitudes were very large, that is, when the reference determinant was not dominant.

Based on an error analysis of the amplitude determining eq. (21), we have introduced an alternative monitoring of the stability based on the ratio of the Euclidean norms of the two subsequent residuals in eq. (21). If this ratio fits into a wide range from 0.1 to 0.95, the overall convergence is guaranteed, while for an optimal computational performance a threshold around 0.5 can be recommended.

It turned out that an approximate update of the t-amplitudes via eq. (22) or via eq. (29) perform equally well, hence the computationally more favorable eq. (22) eventually with the denominator damping as in eq. (29) can be safely recommended.

We have tested the performance of the four-component diagonalization-free SCF approach on eight heavy metal complexes, the largest of which contained more than 80 atoms and about 6000 4-spinor basis functions. In general, a speedup by a factor of two was accomplished for iterations when the diagonalization steps have been replaced by optimizing the Thouless expansion. The gain was achieved without any parallelization, but the present diagonalization-free approach is already suited for running on distributed memory parallel computers employing for instance the ScaLAPACK environment.^[54] This work is in progress in our group.

Acknowledgment

The authors are grateful to Kenneth Ruud for valuable comments on the manuscript.

Appendix A: Proof of the General Validity of Eq. (16) with Arbitrary t-Amplitudes

Theorem. Let us have a basis set \mathcal{B} of one-electron wave functions $\varphi_p, p=1, \dots, n$, and Slater determinant $|\Phi\rangle$, which is composed of one-electron wave functions $\varphi_i, i=1, \dots, N$:

$$|\Phi\rangle = |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle. \quad (\text{A1})$$

Let

$$|\Psi\rangle = e^{T_1} |\Phi\rangle, \quad (\text{A2})$$

where

$$T_1 = \sum_{i,a} t_a^i a_i^a = \sum_{i,a} t_a^i a_a^+ a_i \quad (\text{A3})$$

is a general single excitation operator (t_a^i -s are arbitrary numbers). The one-electron density matrix corresponding to the N -electron wave function $|\Psi\rangle$ in basis \mathcal{B} is then constructed as follows:

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{oo} & \mathbf{D}_{ov} \\ \mathbf{D}_{vo} & \mathbf{D}_{vv} \end{pmatrix}, \quad (\text{A4})$$

where

$$\mathbf{D}_{oo} = (\mathbf{1}_{oo} + \mathbf{t} \mathbf{t}^T)^{-1}, \quad (\text{A5})$$

$$\mathbf{D}_{ov} = \mathbf{D}_{oo} \mathbf{t}, \quad (\text{A6})$$

$$\mathbf{D}_{vo} = \mathbf{t}^T \mathbf{D}_{oo}, \quad (\text{A7})$$

$$\mathbf{D}_{vv} = \mathbf{t}^T \mathbf{D}_{oo} \mathbf{t}, \quad (\text{A8})$$

where \mathbf{t} is a matrix of dimensions $N \times (n-N)$ containing t_a^i as elements.

Proof. We will first prove the following statement:

$$e^{T_1} |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle = |\psi_1 \dots \psi_i \dots \psi_N\rangle, \quad (\text{A9})$$

where

$$|\psi_i\rangle = e^{T_1} |\varphi_i\rangle. \quad (\text{A10})$$

$|\psi_1 \dots \psi_i \dots \psi_N\rangle$ denotes a Slater determinant composed of one-electron wave functions $\psi_i, i=1, \dots, N$. Note that the operator e^{T_1} can be applied to both one-electron and many-electron wave functions. This is because T_1 is just a sum of

operators $a_i^a = a_a^+ a_i$, which are defined for wave functions with arbitrary number of particles.

Let us first rewrite eq. (A10). As

$$a_j^a a_k^b |\varphi_i\rangle = 0 \quad \forall j, k \in \{1, \dots, N\}, a, b \in \{N+1, \dots, n\} \quad (\text{A11})$$

and

$$a_j^a |\varphi_i\rangle = 0 \quad \forall j \neq i, \quad (\text{A12})$$

the following holds

$$\begin{aligned} |\psi_i\rangle &= e^{T_1} |\varphi_i\rangle = \left(1 + T_1 + \frac{T_1^2}{2!} + \dots\right) |\varphi_i\rangle = (1 + T_1) |\varphi_i\rangle \\ &= \left(1 + \sum_a t_a^i a_i^a\right) |\varphi_i\rangle. \end{aligned} \quad (\text{A13})$$

The operator $(1 + \sum_a t_a^i a_i^a)$ can be taken to the front of the Slater determinant:

$$\begin{aligned} |\psi_1 \dots \psi_i \dots \psi_N\rangle &= \left(1 + \sum_a t_a^1 a_1^a\right) \dots \left(1 + \sum_b t_b^i a_i^b\right) \dots \\ &\quad \left(1 + \sum_c t_c^N a_N^c\right) |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle \\ &= (1 + \tau_1) \dots (1 + \tau_i) \dots (1 + \tau_N) |\varphi_1 \dots \varphi_i \dots \varphi_N\rangle \end{aligned}$$

where in the second row, we used the substitution

$$\tau_i = \sum_a t_a^i a_i^a.$$

Note that we can now express T_1 in terms of τ_i :

$$T_1 = \sum_i \tau_i.$$

It remains to show that the operator $(1 + \tau_1) \dots (1 + \tau_i) \dots (1 + \tau_N)$ and the operator e^{T_1} act on the state $|\Phi\rangle$ in the same way. That means, we need to prove the following

$$\begin{aligned} &\left(1 + T_1 + \frac{T_1^2}{2!} + \dots + \frac{T_1^j}{j!} + \dots + \frac{T_1^N}{N!}\right) |\Phi\rangle \\ &= (1 + \tau_1) \dots (1 + \tau_i) \dots (1 + \tau_N) |\Phi\rangle \end{aligned}$$

which is equivalent to

$$\sum_{j=0}^N \frac{(\tau_1 + \dots + \tau_i + \dots + \tau_N)^j}{j!} |\Phi\rangle = (1 + \tau_1) \dots (1 + \tau_i) \dots (1 + \tau_N) |\Phi\rangle. \quad (\text{A14})$$

We have omitted terms with higher powers than N on the left side, because they become 0. (Once all N electrons are excited from their original states, there is nothing left to be excited and the $N+1$ st annihilation operator gets 0). To see eq. (A14) holds, recall that after expanding the left side there will be many zero terms (all those that contain higher than first power of τ_i for any i). Finally, it is easy to check that the

coefficients next to different terms are the same for the right side and the left side. This finishes the proof of eq. (A9) with the definition in eq. (A10).

One last problem remains to be solved before constructing the density matrix. The states ψ_i which we used to construct the new Slater determinant $|\Psi\rangle$ are not orthogonal. However, it is convenient to construct one-particle density matrix in an orthogonal basis. Hence, we would like to find an orthogonal basis $\bar{\psi}_i$ such that its first N elements span the same vector subspace as ψ_1, \dots, ψ_N . Moreover, the density matrix is constructed from a state that is normalized to unity. Thus, we would like the basis $\bar{\psi}_i$ to be orthonormal.

Let us switch to matrix notation. We will write the vectors $\psi_i, i=1, \dots, N$ as columns of a matrix of size $n \times N$ in the φ_i basis[†]

$$\mathbf{W} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ t_{N+1}^1 & t_{N+1}^2 & \cdots & t_{N+1}^N \\ \vdots & \vdots & \vdots & \vdots \\ t_n^1 & t_n^2 & \cdots & t_n^N \end{pmatrix} = \begin{pmatrix} 1_{NN} \\ \mathbf{t}^T \end{pmatrix}$$

and try to find a matrix \mathbf{P} of size $n \times n$ such that the vectors $\bar{\psi}_i$ written in columns of the matrix \mathbf{WP} are orthonormal

$$(\mathbf{WP})^T(\mathbf{WP}) = 1_{NN}$$

In other words, we would like to find N linear combinations of the states $\psi_i, i=1, \dots, N$ that are orthonormal.

It is easy to show that[‡]

$$\mathbf{P} = (\mathbf{W}^T \mathbf{W})^{-1/2} = (1 + \mathbf{t} \mathbf{t}^T)^{-1}$$

meets our requirement:

$$\begin{aligned} (\mathbf{WP})^T(\mathbf{WP}) &= \mathbf{P}^T \mathbf{W}^T \mathbf{W} \mathbf{P} \\ &= (\mathbf{W}^T \mathbf{W})^{-1/2} (\mathbf{W}^T \mathbf{W}) (\mathbf{W}^T \mathbf{W})^{-1/2} \\ &= 1_{NN} \end{aligned}$$

By matrix $\mathbf{S} = \mathbf{M}^{-1/2}$, we mean a matrix such that $(\mathbf{S}^2)^{-1} = \mathbf{M}$. Note that $\mathbf{P} = (\mathbf{W}^T \mathbf{W})^{-1/2}$ always exists for our \mathbf{W} . It can be constructed in the following way: since $\mathbf{W}^T \mathbf{W}$ is symmetric, it is equal to $\mathbf{A}^T \mathbf{B} \mathbf{A}$ where \mathbf{A} is orthonormal and \mathbf{B} is diagonal. Moreover, as it is symmetric and positive definite, all its eigenvalues (which are the diagonal elements of \mathbf{B}) are real and positive. Therefore, we can construct the matrix \mathbf{P} as $\mathbf{A}^T \mathbf{B}^{-1/2} \mathbf{A}$, where $\mathbf{B}^{-1/2}$ is a diagonal matrix, its elements being

[†]Vector $(0, \dots, 0, 1, 0, \dots, 0)^T$ with 1 on the i th place corresponds to φ_i .

[‡]We will omit indices in the identity matrix 1_{NN} , as the dimensions of the matrix are clear from the context and the indices just make the text more difficult to read.

inverses of square roots of diagonal elements of \mathbf{B} . This matrix will satisfy the equation $(\mathbf{P}^2)^{-1} = \mathbf{W}^T \mathbf{W}$ as we wanted.

Now, we just need to construct the one-particle density matrix. As the states $\bar{\psi}_i$ written in columns of \mathbf{WP} are already orthonormal, we obtain the density matrix as follows

$$\mathbf{D} = (\mathbf{WP})(\mathbf{WP})^T = \mathbf{WPP}^T \mathbf{W}^T$$

Substituting for \mathbf{W} and \mathbf{P} leads to

$$\mathbf{D} = \begin{pmatrix} 1 \\ \mathbf{t}^T \end{pmatrix} (1 + \mathbf{t} \mathbf{t}^T)^{-1} \begin{pmatrix} 1 & \mathbf{t} \end{pmatrix} = \begin{pmatrix} (1 + \mathbf{t} \mathbf{t}^T)^{-1} & (1 + \mathbf{t} \mathbf{t}^T)^{-1} \mathbf{t} \\ \mathbf{t}^T (1 + \mathbf{t} \mathbf{t}^T)^{-1} & \mathbf{t}^T (1 + \mathbf{t} \mathbf{t}^T)^{-1} \mathbf{t} \end{pmatrix} \quad (\text{A15})$$

which is what we wanted to prove. ■

Appendix B: Error Analysis of the t-Updating Equations

Using eq. (21) as a starting point, we will here provide a thorough derivation of eq. (22) for the update of the t -amplitudes as proposed in Ref. [39]. This is to provide the reader with insight into the approximations made in updating the t -amplitudes and with an understanding of the need for a good initial guess for the reference MOs.

Let us start from eq. (21). We would like to find the blocks of matrix \mathbf{X} such that this equation is satisfied. We know \mathbf{X}^{old} , which is the value of this matrix from the previous iteration. This matrix provides an initial estimate which is not too far away from the optimal solution. Thus, we will express the new matrix \mathbf{X}^{new} in terms of the old matrix \mathbf{X}^{old} and a correction matrix $\Delta \mathbf{X}$:

$$\mathbf{X}^{\text{new}} = \mathbf{X}^{\text{old}} + \Delta \mathbf{X} \quad (\text{B1})$$

and look for a $\Delta \mathbf{X}$, which satisfies eq. (21). Using eq. (B1) in eq. (21) gives:

$$\begin{aligned} \mathbf{0} &= (\mathbf{1}_{\text{oo}} + \mathbf{X}_{\text{oo}}^{\text{old}}) [\mathbf{F}_{\text{ov}} (\mathbf{1}_{\text{vv}} - \mathbf{X}_{\text{vv}}^{\text{old}}) - \mathbf{F}_{\text{oo}} \mathbf{X}_{\text{ov}}^{\text{old}}] \\ &\quad + \mathbf{X}_{\text{ov}}^{\text{old}} [\mathbf{F}_{\text{vv}} (\mathbf{1}_{\text{vv}} - \mathbf{X}_{\text{vv}}^{\text{old}}) - \mathbf{F}_{\text{vo}} \mathbf{X}_{\text{ov}}^{\text{old}}] \\ &\quad - (\mathbf{1}_{\text{oo}} + \mathbf{X}_{\text{oo}}^{\text{old}}) \mathbf{F}_{\text{ov}} \Delta \mathbf{X}_{\text{vv}} + \Delta \mathbf{X}_{\text{oo}} \mathbf{F}_{\text{ov}} (\mathbf{1}_{\text{vv}} - \mathbf{X}_{\text{vv}}^{\text{old}}) - \Delta \mathbf{X}_{\text{oo}} \mathbf{F}_{\text{ov}} \Delta \mathbf{X}_{\text{vv}} \\ &\quad - (\mathbf{1}_{\text{oo}} + \mathbf{X}_{\text{oo}}^{\text{old}}) \mathbf{F}_{\text{oo}} \Delta \mathbf{X}_{\text{ov}} - \Delta \mathbf{X}_{\text{oo}} \mathbf{F}_{\text{oo}} \mathbf{X}_{\text{ov}}^{\text{old}} - \Delta \mathbf{X}_{\text{oo}} \mathbf{F}_{\text{oo}} \Delta \mathbf{X}_{\text{ov}} \\ &\quad - \mathbf{X}_{\text{ov}}^{\text{old}} \mathbf{F}_{\text{vv}} \Delta \mathbf{X}_{\text{vv}} + \Delta \mathbf{X}_{\text{ov}} \mathbf{F}_{\text{vv}} (\mathbf{1}_{\text{vv}} - \mathbf{X}_{\text{vv}}^{\text{old}}) - \Delta \mathbf{X}_{\text{ov}} \mathbf{F}_{\text{vv}} \Delta \mathbf{X}_{\text{vv}} \\ &\quad - \mathbf{X}_{\text{ov}}^{\text{old}} \mathbf{F}_{\text{vo}} \Delta \mathbf{X}_{\text{ov}} - \Delta \mathbf{X}_{\text{ov}} \mathbf{F}_{\text{vo}} \mathbf{X}_{\text{ov}}^{\text{old}} - \Delta \mathbf{X}_{\text{ov}} \mathbf{F}_{\text{vo}} \Delta \mathbf{X}_{\text{ov}}. \end{aligned} \quad (\text{B2})$$

Note that the expression in the first two lines of eq. (B2) is equal to the ov -block of $\tilde{\mathbf{F}}$ as given by eq. (24) with \mathbf{Q} containing blocks of \mathbf{X}^{old} .

We do not want to solve for $\Delta \mathbf{X}$ directly from eq. (B2). Instead, we want to find

$$\mathbf{t}^{\text{new}} = \mathbf{t}^{\text{old}} + \Delta \mathbf{t} \quad (\text{B3})$$

such that \mathbf{X}^{new} constructed from this \mathbf{t}^{new} satisfies eq. (21). By \mathbf{t}^{old} , we mean \mathbf{t} obtained in the previous iteration. Again, we assume that the change $\Delta \mathbf{t}$ will be small and that \mathbf{t}^{old} is a good approximation for the exact solution for \mathbf{t}^{new} .

We need to express the blocks of $\Delta\mathbf{X}$ in terms of \mathbf{t}^{old} and $\Delta\mathbf{t}$. It is not difficult to show⁵ that for small $\Delta\mathbf{t}$:

$$\Delta\mathbf{X}_{\text{oo}} \approx \Delta\mathbf{x}_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}} \Delta\mathbf{x}_{\text{oo}} + \Delta\mathbf{x}_{\text{oo}} \mathbf{x}_{\text{oo}}^{\text{old}} + \mathbf{x}_{\text{oo}}^{\text{old}} \Delta\mathbf{x}_{\text{oo}} \mathbf{x}_{\text{oo}}^{\text{old}} \\ = -\mathbf{t}^{\text{old}} \Delta\mathbf{t}^{\text{T}} - \Delta\mathbf{t}(\mathbf{t}^{\text{old}})^{\text{T}} + (\text{higher-order terms}) \quad (\text{B4})$$

$$\Delta\mathbf{X}_{\text{ov}} \approx \Delta\mathbf{t} + \Delta\mathbf{X}_{\text{oo}} \mathbf{t}^{\text{old}} + \mathbf{x}_{\text{oo}}^{\text{old}} \Delta\mathbf{t} \\ = \Delta\mathbf{t} + (\text{higher-order terms}) \quad (\text{B5})$$

$$\Delta\mathbf{X}_{\text{vo}} \approx \Delta\mathbf{t}^{\text{T}} + (\mathbf{t}^{\text{old}})^{\text{T}} \Delta\mathbf{X}_{\text{oo}} + \Delta\mathbf{t}^{\text{T}} \mathbf{x}_{\text{oo}}^{\text{old}} \\ = \Delta\mathbf{t}^{\text{T}} + (\text{higher-order terms}) \quad (\text{B6})$$

$$\Delta\mathbf{X}_{\text{vv}} \approx \Delta\mathbf{t}^{\text{T}} \mathbf{t}^{\text{old}} + (\mathbf{t}^{\text{old}})^{\text{T}} \Delta\mathbf{t} + \Delta\mathbf{t}^{\text{T}} \mathbf{x}_{\text{oo}}^{\text{old}} \mathbf{t}^{\text{old}} + (\mathbf{t}^{\text{old}})^{\text{T}} \\ \Delta\mathbf{X}_{\text{oo}} \mathbf{t}^{\text{old}} + \Delta\mathbf{t}^{\text{T}} \mathbf{x}_{\text{oo}}^{\text{old}} \Delta\mathbf{t} \\ = \Delta\mathbf{t}^{\text{T}} \mathbf{t} + \mathbf{t}^{\text{T}} \Delta\mathbf{t} + (\text{higher-order terms}) \quad (\text{B7})$$

where

$$\Delta\mathbf{x}_{\text{oo}} = -\mathbf{t}^{\text{old}} \Delta\mathbf{t}^{\text{T}} - \Delta\mathbf{t}(\mathbf{t}^{\text{old}})^{\text{T}}. \quad (\text{B8})$$

By higher-order terms in the equations above, we mean all terms with these two properties

- they are products of more than two factors, where factors mean \mathbf{t}^{old} , $(\mathbf{t}^{\text{old}})^{\text{T}}$, $\Delta\mathbf{t}$, and $\Delta\mathbf{t}^{\text{T}}$.
- at least one of the product factors is $\Delta\mathbf{t}$ or $\Delta\mathbf{t}^{\text{T}}$.

To justify neglecting these higher-order terms, not only should $\Delta\mathbf{t}$ be small, but also \mathbf{t}^{old} should be significantly smaller than unity. Otherwise, some neglected terms such as $\Delta\mathbf{t}(\mathbf{t}^{\text{old}})^{\text{T}} \mathbf{t}^{\text{old}}$, with elements of \mathbf{t}^{old} of order 1, may be as large as $\Delta\mathbf{t}$ itself. Of course, as $\Delta\mathbf{t}$ and \mathbf{t}^{old} are matrices, there might be lucky cases where large elements from \mathbf{t}^{old} are multiplied by elements from $\Delta\mathbf{t}$ which are approaching zero; but as this is not guaranteed, it is very difficult to accurately estimate or predict the errors caused using our approximations.

To obtain a linear equation for the update of \mathbf{t} , we need to neglect even more terms. In fact, after rewriting eq. (B2) using

⁵The approximate equality for $\Delta\mathbf{X}_{\text{oo}}$ was derived using matrix Taylor expansion of the inverse matrix $(1 - \mathbf{x}_{\text{oo}})^{-1}$. All other formulas follow from this one, as all other blocks of \mathbf{X} are calculated by multiplying $(1_{\text{oo}} + \mathbf{x}_{\text{oo}})$ by \mathbf{t} or \mathbf{t}^{T} either from the left or right.

⁶Except for $\tilde{\mathbf{F}}_{\text{ov}}$, only two terms of the expansion in eq. (B2) will survive: $-(1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}}) \mathbf{F}_{\text{oo}} \Delta\mathbf{X}_{\text{ov}}$ and $\Delta\mathbf{X}_{\text{ov}} \mathbf{F}_{\text{vv}} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{old}})$. We expand $\Delta\mathbf{X}_{\text{ov}}$ using:

$$\mathbf{X}_{\text{ov}}^{\text{new}} = \mathbf{t}^{\text{new}} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{new}}) = (1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{new}}) \mathbf{t}^{\text{new}}.$$

The first equality leads to

$$\Delta\mathbf{X}_{\text{ov}} \approx \Delta\mathbf{t} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{old}}) - \mathbf{t}^{\text{old}} \Delta\mathbf{X}_{\text{vv}} \approx \Delta\mathbf{t} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{old}})$$

and the second one leads to

$$\Delta\mathbf{X}_{\text{ov}} \approx (1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}}) \Delta\mathbf{t} + \Delta\mathbf{X}_{\text{oo}} \mathbf{t}^{\text{old}} \approx (1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}}) \Delta\mathbf{t}.$$

Note that the neglected terms are again of the higher order described previously. Using these approximations leads finally to Eq. (B9).

the aforementioned approximations, we have to neglect all terms that are products of two (or more) factors and that contain at least one $\Delta\mathbf{t}$ or $\Delta\mathbf{t}^{\text{T}}$

$$\mathbf{0} \approx \tilde{\mathbf{F}}_{\text{ov}} - (1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}}) \mathbf{F}_{\text{oo}} (1_{\text{oo}} + \mathbf{x}_{\text{oo}}^{\text{old}}) \Delta\mathbf{t} \\ + \Delta\mathbf{t} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{old}}) \mathbf{F}_{\text{vv}} (1_{\text{vv}} - \mathbf{x}_{\text{vv}}^{\text{old}}) \\ = \tilde{\mathbf{F}}_{\text{ov}} - \tilde{\mathbf{F}}_{\text{oo}} \Delta\mathbf{t} + \Delta\mathbf{t} \tilde{\mathbf{F}}_{\text{vv}} \quad (\text{B9})$$

Finally, if we assume that $\tilde{\mathbf{F}}_{\text{oo}}$ and $\tilde{\mathbf{F}}_{\text{vv}}$ are strongly diagonally dominant, we obtain the first guess of $\Delta\mathbf{t}$ as

$$\Delta t_a^i = \frac{\tilde{F}_a^i}{\tilde{F}_i^i - \tilde{F}_a^a}. \quad (\text{B10})$$

This corresponds to eq. (29) except for the constant Δ in the denominator used to shift orbital energy levels. This technique is also used in conventional SCF procedures especially in quasi-degenerate systems where the orbital energies \tilde{F}_i^i and \tilde{F}_a^a are very close.

Finally, from the structure of \mathbf{Q} [eq. (25)] it follows that the first term in the expansion of $\tilde{\mathbf{F}}$ [eq. (24)] is equal to the Fock matrix in the reference basis. Assuming that this term is dominant, we can replace the orbital energies in eq. (B10) by the reference orbital energies and we obtain the even more approximate estimate of $\Delta\mathbf{t}$ given in eq. (22).

We can now understand the need for good reference MOs. If they are too far from the optimal MOs, the inherent approximations involved in the \mathbf{t} -updating eqs. (22) and (29) may be too severe.

Keywords: diagonalization-free · SCF · computational efficiency · relativistic effects · four-component

How to cite this article: M. Hrdá, T. Kulich, M. Repiský, J. Noga, O. L. Malkina, V. G. Malkin *J. Comput. Chem.* **2014**, *35*, 1725–1737. DOI: 10.1002/jcc.23674



Additional Supporting Information may be found in the online version of this article.

- [1] P. Pyykkö, *Chem. Rev.* **1988**, *88*, 563.
- [2] L. Visscher, *J. Comput. Chem.* **2002**, *23*, 759.
- [3] T. Saue, *ChemPhysChem* **2011**, *12*, 3077.
- [4] L. Belpassi, L. Storchi, H. M. Quiney, F. Tarantelli, *Phys. Chem. Chem. Phys.* **2011**, *13*, 12368.
- [5] T. Fleig, *Chem. Phys.* **2012**, *395*, 2.
- [6] W. Kutzelnigg, *Chem. Phys.* **2012**, *395*, 16.
- [7] W. Liu, *Phys. Chem. Chem. Phys.* **2012**, *14*, 35.
- [8] (a) P. A. M. Dirac, *Proc. R. Soc. London Ser. A* **1928**, *117*, 610; (b) P. A. M. Dirac, *Proc. R. Soc. London Ser. A* **1928**, *118*, 351.
- [9] P. J. C. Aerts, W. C. Nieuwpoort, *Int. J. Quantum Chem. Symp.* **1986**, *19*, 267.
- [10] K. G. Dyall, P. R. Taylor, K. Faegri, Jr., H. Partridge, *J. Chem. Phys.* **1991**, *95*, 2583.
- [11] L. Visscher, O. Visser, H. Aerts, H. Merenga, W. C. Nieuwpoort, *Comput. Phys. Commun.* **1994**, *81*, 120.
- [12] L. Pisani, E. Clementi, *J. Comput. Chem.* **1994**, *15*, 466.

- [13] T. Saue, K. Faegri, Jr., T. Helgaker, O. Gropen, *Mol. Phys.* **1997**, 91, 937.
- [14] H. M. Quiney, H. Skaane, I. P. Grant, *Adv. Quantum Chem.* **1998**, 32, 1.
- [15] T. Yanai, T. Nakajima, Y. Ishikawa, K. Hirao, *J. Chem. Phys.* **2001**, 114, 6526.
- [16] W. Liu, G. Hong, D. Dai, L. Li, M. Dolg, *Theor. Chim. Acta.* **1997**, 96, 75.
- [17] T. Yanai, H. Iikura, T. Nakajima, Y. Ishikawa, K. Hirao, *J. Chem. Phys.* **2001**, 115, 8267.
- [18] T. Saue, T. Helgaker, *J. Comput. Chem.* **2002**, 23, 814.
- [19] H. M. Quiney, P. Belanzoni, *J. Chem. Phys.* **2002**, 117, 5550.
- [20] R. E. Stanton, S. Havriliak, *J. Chem. Phys.* **1984**, 81, 1910.
- [21] W. Kutzelnigg, *Int. J. Quantum Chem.* **1984**, 25, 107.
- [22] S. Komorovsky, M. Repisky, O. L. Malkina, V. G. Malkin, I. M. Ondik, M. Kaupp, *J. Chem. Phys.* **2008**, 128, 104101.
- [23] M. Repisky, S. Komorovsky, O. L. Malkina, V. G. Malkin, *Chem. Phys.* **2009**, 356, 236.
- [24] J. Kussmann, M. Beer, C. Ochsenfeld, *WIREs Comput. Mol. Sci.* **2013**, 3, 614.
- [25] M. Häser, R. Ahlrichs, *J. Comput. Chem.* **1989**, 10, 104.
- [26] L. Visscher, *Theor. Chem. Acc.* **1997**, 98, 68.
- [27] T. Nakajima, K. Hirao, *J. Chem. Phys.* **2004**, 121, 3438.
- [28] S. Varga, E. Engel, W.-D. Sepp, B. Fricke, *Phys. Rev. A* **1999**, 59, 4288.
- [29] (a) L. Belpassi, F. Tarantelli, A. Sgamellotti, H. M. Quiney, *J. Chem. Phys.* **2006**, 124, 124104; (b) L. Belpassi, F. Tarantelli, A. Sgamellotti, H. M. Quiney, *J. Chem. Phys.* **2008**, 128, 124108; (c) L. Storch, L. Belpassi, F. Tarantelli, A. Sgamellotti, H. M. Quiney, *J. Chem. Theory Comput.* **2010**, 6, 384.
- [30] M. Repisky, Development and implementation of efficient relativistic methods for calculations of NMR and EPR parameters, Ph.D. Thesis, Comenius University, Bratislava, 2009.
- [31] M. S. Kelley, T. Shiozaki, *J. Chem. Phys.* **2013**, 138, 204113.
- [32] F. Weigend, *Phys. Chem. Chem. Phys.* **2002**, 4, 4285.
- [33] T. Yanai, T. Nakajima, Y. Ishikawa, K. Hirao, *J. Chem. Phys.* **2002**, 116, 10122.
- [34] I. P. Grant, H. M. Quiney, In *Relativistic Self-Consistent Fields in Relativistic Electronic Structure Theory. Part 1. Fundamentals*; P. Schwerdtfeger, Ed.; Elsevier: Amsterdam, **2002**.
- [35] M. Repisky, *InterRest 2.0, An integral program for relativistic quantum chemistry*, 2013.
- [36] T. Shiozaki, *J. Chem. Phys.* **2013**, 138, 111101.
- [37] L. Storch, S. Rampino, L. Belpassi, F. Tarantelli, H. M. Quiney, *J. Chem. Theory Comput.* **2013**, 9, 5356.
- [38] M. Repisky, S. Komorovsky, S. Knecht, K. Ruud, *submitted to Journal of Chemical Theory and Computation*.
- [39] J. Noga, J. Šimunek, *J. Chem. Theory Comput.* **2010**, 6, 2706.
- [40] A. Baldes, W. Klopper, J. Šimunek, J. Noga, F. Weigend, *J. Comput. Chem.* **2011**, 32, 3129.
- [41] M. Repisky, S. Komorovsky, V. G. Malkin, O. L. Malkina, M. Kaupp, K. Ruud, With contributions from R. Bast, U. Ekstrom, S. Knecht, I. Malkin Ondik, E. Malkin, *Relativistic Spectroscopy DFT program, ReSpect, version 3.3.0*, **2013**, Available at <http://rel-qchem.sav.sk>. Last accessed on June 28 2014.
- [42] P. R. Surjan, A. Szabados, Challenges and Advances in Computational Chemistry and Physics, Vol. 13; R. Zalesny, M. G. Papadopoulos, P. G. Mezey, J. Leszczynski, Eds.; Springer: Dordrecht, 2011; pp. 83–95.
- [43] W. Z. Liang, M. Head-Gordon, *J. Chem. Phys.* **2004**, 120, 10379.
- [44] S. Høst, J. Olsen, B. Jansík, L. Thørgensen, P. Jørgensen, T. Helgaker, *J. Chem. Phys.* **2008**, 129, 124106.
- [45] D. J. Thouless, *Nucl. Phys.* **1960**, 21, 225.
- [46] J. Šimunek, J. Noga, *AlP Conf. Proc.* **2012**, 1504, 143.
- [47] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, 56, 2257.
- [48] (a) W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.* **1969**, 51, 2657; (b) W. J. Pietro, B. A. Levy, W. J. Hehre, R. F. Stewart, *J. Am. Chem. Soc.* **1980**, 102, 2225.
- [49] (a) F. Jensen, *J. Chem. Phys.* **2001**, 115, 9113; (b) F. Jensen, *J. Chem. Phys.* **2002**, 116, 3502.
- [50] A. D. Becke, *Phys. Rev. A* **1988**, 38, 3098.
- [51] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, 37, 785.
- [52] (a) K. G. Dyall, *Theor. Chem. Acc.* **2006**, 115, 441; (b) K. G. Dyall, *Theor. Chem. Acc.* **2007**, 117, 483; (c) K. G. Dyall, A. S. P. Gomes, *Theor. Chem. Acc.* **2009**, 125, 97; (d) K. G. Dyall, *Theor. Chem. Acc.* **2007**, 117, 491; (e) Basis sets, Dirac web site, Available at <http://dirac.chem.sdu.dk>. Last accessed on June 28 2014.
- [53] (a) T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, 90, 1007; (b) D. E. Woon, T. H. Dunning, Jr., *J. Chem. Phys.* **1993**, 98, 1358.
- [54] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, R. C. Whaley, *ScaLAPACK Users' Guide*; Society for Industrial and Applied Mathematics: Philadelphia, PA, 1997.

Received: 23 April 2014
Revised: 9 June 2014
Accepted: 10 June 2014
Published online on 4 July 2014