Development and application of the analytical energy gradient for the normalized elimination of the small component method

Wenli Zou,[1] Michael Filatov,[2] and Dieter Cremer[1,a)]
[1]*Department of Chemistry, Southern Methodist University, 3215 Daniel Ave, Dallas, Texas 75275-0314, USA*
[2]*University of Groningen, Nijenborgh 4, NL-9747AG Groningen, The Netherlands*

(Received 17 April 2011; accepted 4 June 2011; published online 30 June 2011)

The analytical energy gradient of the normalized elimination of the small component (NESC) method is derived for the first time and implemented for the routine calculation of NESC geometries and other first order molecular properties. Essential for the derivation is the correct calculation of the transformation matrix **U** relating the small component to the pseudolarge component of the wavefunction. The exact form of $\partial \mathbf{U}/\partial\lambda$ is derived and its contribution to the analytical energy gradient is investigated. The influence of a finite nucleus model and that of the picture change is determined. Different ways of speeding up the calculation of the NESC gradient are tested. It is shown that first order properties can routinely be calculated in combination with Hartree-Fock, density functional theory (DFT), coupled cluster theory, or any electron correlation corrected quantum chemical method, provided the NESC Hamiltonian is determined in an efficient, but nevertheless accurate way. The general applicability of the analytical NESC gradient is demonstrated by benchmark calculations for NESC/CCSD (coupled cluster with all single and double excitation) and NESC/DFT involving up to 800 basis functions. © *2011 American Institute of Physics*. [doi:10.1063/1.3603454]

I. INTRODUCTION

The usefulness of a new quantum chemical method is measured by its accuracy and general applicability. General applicability implies the possibility of routinely calculating molecular properties, such as geometries, vibrational frequencies, electric or magnetic quantities, excitation energies, etc. A large number of molecular properties needed by chemists to analyze and characterize structure, stability, and reactivity of a chemical compound are response properties, which can be effectively calculated with the help of analytical energy derivatives. Accordingly, the usefulness of a quantum mechanical method increases substantially when its applicability range is extended by the introduction of analytical energy derivatives.

The calculation of exact quasi-relativistic energies identical in their values to the results of four-component (4c) calculations based on the Dirac equation[1,2] can be considered as a breakthrough in relativistic quantum chemistry.[3–12] This progress was started with Dyall's work on the normalized elimination of the small component (NESC) method,[3,4] which provided for the first time exact 4c-energies at the one (1c)- or two-component (2c) level. The basis for this breakthrough was a revised strategy to solve the problem of transforming the 4c relativistic problem into a 2c or 1c quasi-relativistic one: Rather than deriving an approximate 2c-Hamiltonian at the operator level (operator-driven approaches), Dyall expressed the Dirac Hamiltonian in matrix form using a finite basis set and then, by using matrix algebra, derived a simplified quasi-relativistic Hamiltonian (matrix-driven approach).[3,4] It needed some time to cast NESC in a generally applicable form, which was first accomplished by Filatov and Dyall.[13] Systematic application of NESC was carried out by Cremer, Filatov, and Kraka,[14,15] Kraka and Cremer,[16] and Filatov *et al.*[17–21]

Dyall's work triggered directly or indirectly a number of other important developments in the area of exact quasi-relativistic methods. Reiher, Wolf, and co-workers, following early work by Hess and others on Douglas Kroll theory,[22–28] developed infinite order Douglas-Kroll-Hess theory,[5–7,29] which also can provide exact 4c energies at the 1c or 2c level of theory. Early work by Barysz, Sadlej, and Snijders[30] on the infinite order two-component approach[31–35] led to a matrix-driven formulation of this method and exact quasi-relativistic energies.[8,9,36] Kutzelnigg and Liu carried out a fundamental analysis of the possibilities of transforming the Dirac Hamiltonian into a quasi-relativistic Hamiltonian. They set up conditions and algorithms for obtaining exact quasi-relativistic (X2C or XQC) methods.[10–12] The pioneering work of Dyall on the NESC method was clarified by Filatov[37] and Kutzelnigg and Liu.[38]

In this work, we make the next step in the development of exact quasi-relativistic methods by presenting for the first time the analytical energy gradient of the NESC method. For this purpose, our work is organized in the following way. In Sec. II, we summarized the basics of calculating NESC energies. In this connection, we will shortly point out some new ways of effectively solving the NESC equations. In Sec. III, the analytical energy gradient of the NESC method is derived where special emphasis is laid on the correct analytical form of the transformation matrix **U** relating the small component to the pseudolarge component of the wavefunction. In Sec. IV, we describe the computational techniques used in this work, and finally, in Sec. V, some benchmark calculations are

a)Author to whom correspondence should be addressed. Electronic mail: dieter.cremer@gmail.com.

presented that demonstrate the accuracy and general applicability of the NESC gradient approach presented in this work.

II. CALCULATION OF THE NESC ENERGY

The NESC method requires the solving of Eq. (1) (Refs. 3, 13, and 37) thus providing the electronic (positive-energy) solutions of the Dirac equation:[1]

$$\tilde{\mathbf{L}}\mathbf{A} = \tilde{\mathbf{S}}\mathbf{A}\boldsymbol{\varepsilon}. \tag{1}$$

In Eq. (1), $\tilde{\mathbf{L}}$ represents the NESC Hamiltonian, matrix $\mathbf{A}$ collects the NESC eigenvectors $\mathbf{a}_\mu$ with $\mu = 1, \ldots, M$ (M is the number of basis functions of basis χ), and $\boldsymbol{\varepsilon}$ contains on its diagonal the eigenvalues of the NESC Hamiltonian. The eigenvectors $\mathbf{a}_\mu$ are orthonormalized on the relativistic metric $\tilde{\mathbf{S}}$:

$$\tilde{\mathbf{S}} = \mathbf{S} + \frac{1}{2mc^2}\mathbf{U}^\dagger\mathbf{T}\mathbf{U}, \tag{2}$$

according to

$$\mathbf{A}^\dagger\tilde{\mathbf{S}}\mathbf{A} = \mathbf{I}. \tag{3}$$

The NESC Hamiltonian $\tilde{\mathbf{L}}$ is given by

$$\tilde{\mathbf{L}} = \mathbf{T}\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} - \mathbf{U}^\dagger(\mathbf{T} - \mathbf{W})\mathbf{U} + \mathbf{V}. \tag{4}$$

In Eqs. (1), (2), and (4), $\mathbf{S}$, $\mathbf{T}$, and $\mathbf{V}$ are the matrices of the nonrelativistic overlap, kinetic energy, and potential energy operators, and $\mathbf{W}$ is the matrix of the operator $1/(4m^2c^2)(\boldsymbol{\sigma}\cdot\mathbf{p})V(\mathbf{r})(\boldsymbol{\sigma}\cdot\mathbf{p})$ in the basis of the atomic orbitals $\chi_\mu(\mathbf{r})$,[3] where $\boldsymbol{\sigma}$ and $\mathbf{p}$ are the vector of the Pauli matrices and the linear momentum operator, respectively. At the scalar-relativistic level used throughout this work the latter expression simplifies to $1/(4m^2c^2)\nabla V(\mathbf{r})\cdot\nabla$. The matrix $\mathbf{U}$ is associated with the operator eliminating the small component of the electronic (positive energy) relativistic wave function, i.e., $\mathbf{U}$ connects the matrix of eigenvectors for the large component, $\mathbf{A}$, with the matrix of eigenvectors of the pseudolarge component, $\mathbf{B}$, in the modified Dirac wave function[3,4] according to

$$\mathbf{B} = \mathbf{U}\mathbf{A}. \tag{5}$$

As a suitable starting guess for matrix $\mathbf{U}$ the infinite order regular approximation (*IORA*) equation (6) is solved:

$$\tilde{\mathbf{L}}^{IORA}\mathbf{A}^{IORA} = \tilde{\mathbf{S}}^{IORA}\mathbf{A}^{IORA}\boldsymbol{\varepsilon}^{IORA}, \tag{6}$$

where the Hamiltonian $\tilde{\mathbf{L}}^{IORA}$ and the metric $\tilde{\mathbf{S}}^{IORA}$ are obtained using Eq. (7) in Eqs. (4) and (2):[39]

$$\mathbf{U}^{IORA} = (\mathbf{T} - \mathbf{W})^{-1}\mathbf{T}. \tag{7}$$

Once a starting value is obtained for $\mathbf{U}$, Eqs. (8) and (9),

$$\mathbf{T}\mathbf{U} = \mathbf{S}\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}} - \mathbf{V}, \tag{8}$$

$$\tilde{\mathbf{L}} = (\mathbf{T}\mathbf{U}) + (\mathbf{T}\mathbf{U})^\dagger - (\mathbf{T}\mathbf{U})^\dagger(\mathbf{T}^{-1} - \mathbf{T}^{-1}\mathbf{W}\mathbf{T}^{-1})(\mathbf{T}\mathbf{U}) + \mathbf{V}, \tag{9}$$

are used to iteratively solve for the product $\mathbf{TU}$ and the NESC Hamiltonian matrix $\tilde{\mathbf{L}}$ in the form of a fixed-point iteration for a nonlinear problem:

$$\mathbf{T}\mathbf{U}^{(n)} = \mathbf{F}(\mathbf{T}\mathbf{U}^{(n-1)}). \tag{10}$$

It turns out to be computationally feasible to obtain the solution applying a damped fixed-point iteration technique often employed to solve stiff initial value problems.[40] If the direct fixed-point iteration of Eq. (10) does not converge, one can introduce a damping factor α to stabilize the iteration:

$$\mathbf{T}\mathbf{U}^{(n)} = \mathbf{F}(\mathbf{T}\mathbf{U}^{(n-1)}) - \alpha(\mathbf{F}(\mathbf{T}\mathbf{U}^{(n-1)}) - \mathbf{T}\mathbf{U}^{(n-1)}). \tag{11}$$

In practice, static rather dynamic damping is sufficient to lead to the convergence of the NESC equations (4) and (8). When convergence of $\mathbf{TU}$ is achieved, matrix $\mathbf{U}$ is calculated with the help of the known $\mathbf{T}$ ("iterative TU" method).

Thus, the flowchart of the iterative solution of the NESC equation (1) is organized as in the following scheme:

Step A: $i = 0$,

$$(\mathbf{T}\mathbf{U}^{(0)}) = \mathbf{T}\mathbf{U}^{IORA} = \mathbf{T}(\mathbf{T} - \mathbf{W})^{-1}\mathbf{T};$$

Step B: $i = i + 1$,

$$\ldots \quad (\mathbf{T}\mathbf{U}^{(i)}) = \mathbf{S}(\tilde{\mathbf{S}}^{(i-1)})^{-1}\tilde{\mathbf{L}}^{(i-1)} - \mathbf{V}, \quad \text{(see Eq. (11) for damping)},$$

$$\ldots \quad \tilde{\mathbf{L}}^{(i)} = (\mathbf{T}\mathbf{U}^{(i)}) + (\mathbf{T}\mathbf{U}^{(i)})^\dagger - (\mathbf{T}\mathbf{U}^{(i)})^\dagger(\mathbf{T}^{-1} - \mathbf{T}^{-1}\mathbf{W}\mathbf{T}^{-1})(\mathbf{T}\mathbf{U}^{(i)}) + \mathbf{V},$$

$$\ldots \quad \tilde{\mathbf{S}}^{(i)} = \mathbf{S} + (1/2mc^2)(\mathbf{T}\mathbf{U}^{(i)})^\dagger\mathbf{T}^{-1}(\mathbf{T}\mathbf{U}^{(i)}),$$

$$\ldots \quad \Delta^{(i)} = ||\tilde{\mathbf{L}}_{kk}^{(i)} - \tilde{\mathbf{L}}_{kk}^{(i-1)}||;$$

Conv.: if $\Delta^{(i)} \leq \delta$, exit. Else, go to B.

In the case of geometry optimizations, PES scans, or any other calculation that involves a restart from a previous NESC calculation, the initial $(\mathbf{TU}^{(0)})$ is best obtained from the previously stored $\mathbf{U}$ matrix, which can usually save one to two thirds of the computational cost for solving the NESC equations.

This solution scheme works for all normal basis sets (exponent of steep basis functions χ_μ not larger than 10^9) in connection with the strategy "first-diagonalize-then-contract," which implies that the NESC equations are solved for the uncontracted rather than the contracted basis set, and once the NESC Hamiltonian has been obtained, a transformation to the contracted basis set is carried out. This approach does not require much additional computer time; however, it has a number of advantages with regard to the previously used strategy "first-contract-then-diagonalize:"[13] (i) The NESC iterations converge even when using basis functions with very large exponents ("steep functions"); (ii) Generally, convergence is accelerated; (iii) The convergence of the NESC iterations can be better controlled; (iv) The atomic many-electron total energies are within less than 1 hartree from the energies obtained when employing the corresponding uncontracted basis set.

Equation (1) provides the exact electronic solutions of the 4-component one-electron problem. For the calculation of a many-electron system, the use of a one-electron spin-scalar approximation as suggested by Dyall[4] is adopted in the present work. Within this approximation, the NESC one-electron Hamiltonian equation (4) is renormalized on the non-relativistic metric (12) (Ref. 4) and then used with the non-relativistic many-electron equations, such as the Hartree-Fock or Kohn-Sham equations:

$$\mathbf{H}_{1-e} = \mathbf{G}^\dagger \tilde{\mathbf{L}} \mathbf{G}. \tag{12}$$

Dyall suggested to use Eq. (13) for the renormalization matrix,[4]

$$\mathbf{G} = \tilde{\mathbf{S}}^{-1/2} \mathbf{S}^{1/2}, \tag{13}$$

which, however, may lead to incorrect transformation properties of the renormalized $1-e$ Hamiltonian. In non-relativistic as well as in relativistic theory, the Hamiltonian (and other operator matrices) should transform under a linear transformation of the basis set $\boldsymbol{\chi}' = \boldsymbol{\chi}\mathbf{O}$ as in Eq. (14):

$$\mathbf{H}' = \mathbf{O}^\dagger \mathbf{H} \mathbf{O}. \tag{14}$$

With the use of conventional non-orthogonal basis sets, Eq. (13) does not guarantee this property. Therefore, it was suggested by Liu and Peng[41] to employ (among others) the following transformation matrix:

$$\mathbf{G} = \mathbf{S}^{-1/2} (\mathbf{S}^{1/2} \tilde{\mathbf{S}}^{-1} \mathbf{S}^{1/2})^{1/2} \mathbf{S}^{1/2}. \tag{15}$$

Equation (15) for $\mathbf{G}$ is adopted in this work because it operates with square roots of symmetric matrices, which considerably simplifies derivation of the analytic energy gradient of the renormalized NESC Hamiltonian equation (12).

With the use of the renormalized NESC Hamiltonian equation (12), the total electronic energy is determined, at the Hartree-Fock level, by Eq. (16),

$$E = tr\mathbf{P}\mathbf{H}_{1-e} + \tfrac{1}{2} tr\mathbf{P}(\mathbf{J} - \mathbf{K}), \tag{16}$$

where $\mathbf{J}$ and $\mathbf{K}$ are the Coulomb and the exchange parts of the Fock operator and $\mathbf{P}$ is the density matrix calculated as $\mathbf{P} = \mathbf{C}^F \mathbf{n} (\mathbf{C}^F)^\dagger$ ($\mathbf{C}^F$ collects the eigenvectors of the Fock operator and $\mathbf{n}$ is the diagonal matrix of the orbital occupation numbers).

III. THE ANALYTICAL ENERGY GRADIENT FOR NESC

The first derivative of the total energy equation (16) with respect to the nuclear coordinate λ is given by Eq. (17),[42]

$$\frac{\partial E}{\partial \lambda} = tr\mathbf{P}\left(\frac{\partial \mathbf{H}_{1-e}}{\partial \lambda}\right) + tr\boldsymbol{\Omega}\left(\frac{\partial \mathbf{S}}{\partial \lambda}\right) + \frac{1}{2} tr\mathbf{P}\frac{\partial'}{\partial \lambda}(\mathbf{J} - \mathbf{K}), \tag{17}$$

where $\boldsymbol{\Omega} = -\mathbf{C}^F \mathbf{n} \boldsymbol{\varepsilon} (\mathbf{C}^F)^\dagger$ is the energy-weighted density matrix ($\boldsymbol{\varepsilon}$ collects the orbital energies on its diagonal) and the prime of $\partial'/\partial\lambda$ implies that only the two-electron integrals rather than the density matrix have to be differentiated.[42]

The first term on the rhs of Eq. (17) can be written as

$$\frac{\partial E_{1-e}}{\partial \lambda} = tr\mathbf{P}\mathbf{G}^\dagger \frac{\partial \tilde{\mathbf{L}}}{\partial \lambda}\mathbf{G} + tr\mathbf{P}\frac{\partial \mathbf{G}^\dagger}{\partial \lambda}\tilde{\mathbf{L}}\mathbf{G} + tr\mathbf{P}\mathbf{G}^\dagger \tilde{\mathbf{L}}\frac{\partial \mathbf{G}}{\partial \lambda} \tag{18}$$

$$= tr\mathbf{G}\mathbf{P}\mathbf{G}^\dagger \frac{\partial \tilde{\mathbf{L}}}{\partial \lambda} + tr\tilde{\mathbf{L}}\mathbf{G}\mathbf{P}\frac{\partial \mathbf{G}^\dagger}{\partial \lambda} + tr\mathbf{P}\mathbf{G}^\dagger \tilde{\mathbf{L}}\frac{\partial \mathbf{G}}{\partial \lambda} \tag{19}$$

$$= tr\tilde{\mathbf{P}}\frac{\partial \tilde{\mathbf{L}}}{\partial \lambda} + tr\mathbf{D}\frac{\partial \mathbf{G}^\dagger}{\partial \lambda} + tr\mathbf{D}^\dagger \frac{\partial \mathbf{G}}{\partial \lambda}, \tag{20}$$

where $\mathbf{G}$ is given in Eqs. (13) and (15), $\tilde{\mathbf{P}} = \mathbf{G}\mathbf{P}\mathbf{G}^\dagger$; also $\mathbf{D} = \tilde{\mathbf{L}}\mathbf{G}\mathbf{P}$ and the cyclic property of a trace is used.

Differentiating Eq. (4) with respect to the nuclear coordinate λ yields Eq. (21):

$$\frac{\partial \tilde{\mathbf{L}}}{\partial \lambda} = \frac{\partial \mathbf{T}}{\partial \lambda}\mathbf{U} + \mathbf{U}^\dagger \frac{\partial \mathbf{T}}{\partial \lambda} - \mathbf{U}^\dagger \left(\frac{\partial \mathbf{T}}{\partial \lambda} - \frac{\partial \mathbf{W}}{\partial \lambda}\right)\mathbf{U} + \frac{\partial \mathbf{V}}{\partial \lambda} + \frac{\partial \mathbf{U}^\dagger}{\partial \lambda}(\mathbf{T} - (\mathbf{T} - \mathbf{W})\mathbf{U}) + (\mathbf{T} - \mathbf{U}^\dagger(\mathbf{T} - \mathbf{W}))\frac{\partial \mathbf{U}}{\partial \lambda}. \tag{21}$$

Before discussing the contributions to $\partial\tilde{\mathbf{L}}/\partial\lambda$, we will first consider the derivative of the renormalization matrix $\mathbf{G}$.

A. Derivatives of the renormalization matrix G

Using Eq. (13) for the renormalization matrix $\mathbf{G}$, one obtains the following derivative terms:

$$\frac{\partial \mathbf{G}}{\partial \lambda} = -\tilde{\mathbf{S}}^{-1/2}\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda}\mathbf{G} + \tilde{\mathbf{S}}^{-1/2}\frac{\partial \mathbf{S}^{1/2}}{\partial \lambda}, \tag{22}$$

which yield Eqs. (23) and (24):

$$tr\mathbf{D}\frac{\partial \mathbf{G}^{\dagger}}{\partial \lambda} + tr\mathbf{D}^{\dagger}\frac{\partial \mathbf{G}}{\partial \lambda}$$
$$= -tr\mathbf{D}\mathbf{G}^{\dagger}\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda}\tilde{\mathbf{S}}^{-1/2} + tr\mathbf{D}\frac{\partial \mathbf{S}^{1/2}}{\partial \lambda}\tilde{\mathbf{S}}^{-1/2}$$
$$-tr\mathbf{D}^{\dagger}\tilde{\mathbf{S}}^{-1/2}\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda}\mathbf{G} + tr\mathbf{D}^{\dagger}\tilde{\mathbf{S}}^{-1/2}\frac{\partial \mathbf{S}^{1/2}}{\partial \lambda} \quad (23)$$

$$= -tr(\tilde{\mathbf{S}}^{-1/2}\mathbf{D}\mathbf{G}^{\dagger} + \mathbf{G}\mathbf{D}^{\dagger}\tilde{\mathbf{S}}^{-1/2})\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda}$$
$$+tr(\tilde{\mathbf{S}}^{-1/2}\mathbf{D} + \mathbf{D}^{\dagger}\tilde{\mathbf{S}}^{-1/2})\frac{\partial \mathbf{S}^{1/2}}{\partial \lambda} \quad (24)$$

for the last two terms in Eq. (20).

The derivatives of the square-root of a symmetric positive definite matrix **M** can be obtained utilizing Eq. (25),[43]

$$\left(\mathbf{C}^{\dagger}\frac{\partial \mathbf{M}^{1/2}}{\partial \lambda}\mathbf{C}\right)_{ij} = \left(m_i^{1/2} + m_j^{1/2}\right)^{-1}\left(\mathbf{C}^{\dagger}\frac{\partial \mathbf{M}}{\partial \lambda}\mathbf{C}\right)_{ij}, \quad (25)$$

where **C** contains the eigenvectors and **m** are the eigenvalues of **M**. The derivatives of the overlap integrals are routinely available in the quantum chemical program codes and the derivatives of the relativistic metric $\tilde{\mathbf{S}}$ can be obtained using Eq. (26):

$$\frac{\partial \tilde{\mathbf{S}}}{\partial \lambda} = \frac{\partial \mathbf{S}}{\partial \lambda} + \frac{1}{2mc^2}\mathbf{U}^{\dagger}\frac{\partial \mathbf{T}}{\partial \lambda}\mathbf{U}$$
$$+\frac{1}{2mc^2}\left(\frac{\partial \mathbf{U}^{\dagger}}{\partial \lambda}\mathbf{T}\mathbf{U} + \mathbf{U}^{\dagger}\mathbf{T}\frac{\partial \mathbf{U}}{\partial \lambda}\right). \quad (26)$$

The third term in Eq. (26) depends on the derivatives $\partial\mathbf{U}/\partial\lambda$, which will be discussed in the following (see Eq. (35)).

The derivatives of $\partial\mathbf{M}^{1/2}/\partial\lambda$ can be obtained by multiplying the left hand side of Eq. (25) with the eigenvectors **C** from the left and with $\mathbf{C}^{\dagger}$ from the right as in Eq. (27):[44]

$$\left(\frac{\partial \mathbf{M}^{1/2}}{\partial \lambda}\right)_{mn} = \sum_i\sum_j\sum_k\sum_l C_{mi}\left(m_i^{1/2} + m_j^{1/2}\right)^{-1}$$
$$\times C_{ik}^{\dagger}\left(\frac{\partial \mathbf{M}}{\partial \lambda}\right)_{kl}C_{lj}C_{jn}^{\dagger}. \quad (27)$$

Using Eq. (27) and denoting the eigenvector matrices of $\tilde{\mathbf{S}}$ and **S** as $\tilde{\mathbf{C}}$ and **C** and the two terms of Eqs. (24) given in parentheses as $\tilde{\mathbf{X}}$ and **X**, respectively, the first part of Eq. (24) becomes

$$tr\tilde{\mathbf{X}}\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda} = \sum_m\sum_n \tilde{X}_{nm}\left(\frac{\partial \tilde{\mathbf{S}}^{1/2}}{\partial \lambda}\right)_{mn} = \sum_m\sum_n\sum_k\sum_l\sum_i\sum_j \tilde{X}_{nm}\tilde{C}_{mi}\left(\tilde{s}_i^{1/2} + \tilde{s}_j^{1/2}\right)^{-1}\tilde{C}_{ik}^{\dagger}\left(\frac{\partial \tilde{S}}{\partial \lambda}\right)_{kl}\tilde{C}_{lj}\tilde{C}_{jn}^{\dagger}$$
$$= \sum_k\sum_l\sum_i\sum_j \left(\tilde{s}_i^{1/2} + \tilde{s}_j^{1/2}\right)^{-1}\tilde{C}_{ik}^{\dagger}\left(\frac{\partial \tilde{S}}{\partial \lambda}\right)_{kl}\tilde{C}_{lj}\sum_m\sum_n \tilde{C}_{jn}^{\dagger}\tilde{X}_{nm}\tilde{C}_{mi}$$
$$= \sum_k\sum_l\left(\frac{\partial \tilde{S}}{\partial \lambda}\right)_{kl}\sum_i\sum_j \left(\tilde{s}_i^{1/2} + \tilde{s}_j^{1/2}\right)^{-1}\tilde{C}_{lj}\tilde{Y}_{ji}\tilde{C}_{ik}^{\dagger} = \sum_k\sum_l\left(\frac{\partial \tilde{S}}{\partial \lambda}\right)_{kl}\tilde{Z}_{lk} = tr\tilde{\mathbf{Z}}\frac{\partial \tilde{\mathbf{S}}}{\partial \lambda}, \quad (28)$$

where two new matrices are introduced: $\tilde{\mathbf{Y}} = \tilde{\mathbf{C}}^{\dagger}\tilde{\mathbf{X}}\tilde{\mathbf{C}}$ and $\tilde{\mathbf{Z}}$ with elements $\tilde{Z}_{ij} = \sum_{k,l}\tilde{C}_{ik}\tilde{Y}_{kl}\tilde{C}_{lj}^{\dagger}(\tilde{s}_k^{1/2} + \tilde{s}_l^{1/2})^{-1}$. Similar equations can be derived for matrix **X**. Thus, the sum of the last two terms in Eq. (20) can be evaluated according to Eq. (29):

$$tr\mathbf{D}\frac{\partial \mathbf{G}^{\dagger}}{\partial \lambda} + tr\mathbf{D}^{\dagger}\frac{\partial \mathbf{G}}{\partial \lambda} = -tr\tilde{\mathbf{Z}}\frac{\partial \tilde{\mathbf{S}}}{\partial \lambda} + tr\mathbf{Z}\frac{\partial \mathbf{S}}{\partial \lambda} = tr(\mathbf{Z} - \tilde{\mathbf{Z}})\frac{\partial \mathbf{S}}{\partial \lambda} - \frac{1}{2mc^2}tr(\mathbf{U}\tilde{\mathbf{Z}}\mathbf{U}^{\dagger})\frac{\partial \mathbf{T}}{\partial \lambda} - \frac{1}{2mc^2}tr(\mathbf{T}\mathbf{U}\tilde{\mathbf{Z}})\frac{\partial \mathbf{U}^{\dagger}}{\partial \lambda} - \frac{1}{2mc^2}tr(\tilde{\mathbf{Z}}\mathbf{U}^{\dagger}\mathbf{T})\frac{\partial \mathbf{U}}{\partial \lambda}, \quad (29)$$

where $\mathbf{Y} = \mathbf{C}^{\dagger}\mathbf{X}\mathbf{C}$ and matrix **Z** has elements $Z_{ij} = \sum_{k,l}C_{ik}Y_{kl}C_{lj}^{\dagger}(s_k^{1/2} + s_l^{1/2})^{-1}$.

If **G** is defined by Eq. (15), the gradient contribution of $\partial\mathbf{G}/\partial\lambda$ will become

$$tr\left(\mathbf{D}\frac{\partial \mathbf{G}^{\dagger}}{\partial \lambda}\right) + tr\left(\mathbf{D}^{\dagger}\frac{\partial \mathbf{G}}{\partial \lambda}\right) = tr\mathbf{D}_0\frac{\partial}{\partial \lambda}(\mathbf{S}^{1/2}) + tr\mathbf{D}_1\frac{\partial}{\partial \lambda}(\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1}\mathbf{S}^{1/2})^{1/2}$$
$$= tr\mathbf{D}_{0Z}\frac{\partial \mathbf{S}}{\partial \lambda} + tr\mathbf{D}_{1Z}\frac{\partial}{\partial \lambda}(\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1}\mathbf{S}^{1/2}) = tr\mathbf{D}_{0Z}\frac{\partial \mathbf{S}}{\partial \lambda} + tr\mathbf{D}_2\frac{\partial}{\partial \lambda}(\mathbf{S}^{1/2}) - tr\mathbf{D}_3\frac{\partial}{\partial \lambda}(\tilde{\mathbf{S}})$$
$$= tr\mathbf{D}_{0Z}\frac{\partial \mathbf{S}}{\partial \lambda} + tr\mathbf{D}_{2Z}\frac{\partial \mathbf{S}}{\partial \lambda} - tr\mathbf{D}_3\frac{\partial \mathbf{S}}{\partial \lambda} - \frac{1}{2mc^2}tr(\mathbf{U}\mathbf{D}_3\mathbf{U}^{\dagger})\frac{\partial \mathbf{T}}{\partial \lambda} - \frac{1}{2mc^2}tr\left(\mathbf{T}\mathbf{U}\mathbf{D}_3\frac{\partial \mathbf{U}^{\dagger}}{\partial \lambda} + \mathbf{D}_3\mathbf{U}^{\dagger}\mathbf{T}\frac{\partial \mathbf{U}}{\partial \lambda}\right)$$
$$= tr\left(\mathbf{D}_{0Z} + \mathbf{D}_{2Z} - \mathbf{D}_3\right)\frac{\partial \mathbf{S}}{\partial \lambda} - \frac{1}{2mc^2}tr(\mathbf{U}\mathbf{D}_3\mathbf{U}^{\dagger})\frac{\partial \mathbf{T}}{\partial \lambda} - \frac{1}{2mc^2}tr\left(\mathbf{T}\mathbf{U}\mathbf{D}_3\frac{\partial \mathbf{U}^{\dagger}}{\partial \lambda} + \mathbf{D}_3\mathbf{U}^{\dagger}\mathbf{T}\frac{\partial \mathbf{U}}{\partial \lambda}\right), \quad (30)$$

where

$$\mathbf{D}_0 = (\mathbf{D}^\dagger\mathbf{G} - \mathbf{G}\mathbf{D}^\dagger)\mathbf{S}^{-1/2} + \mathbf{S}^{-1/2} \times (\mathbf{G}^\dagger\mathbf{D} - \mathbf{D}\mathbf{G}^\dagger), \tag{31}$$

$$\mathbf{D}_1 = \mathbf{S}^{1/2}\mathbf{D}^\dagger\mathbf{S}^{-1/2} + \mathbf{S}^{-1/2}\mathbf{D}\mathbf{S}^{1/2}, \tag{32}$$

$$\mathbf{D}_2 = \mathbf{D}_{1Z}\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1} + \tilde{\mathbf{S}}^{-1}\mathbf{S}^{1/2}\mathbf{D}_{1Z}, \tag{33}$$

$$\mathbf{D}_3 = \tilde{\mathbf{S}}^{-1}\mathbf{S}^{1/2}\mathbf{D}_{1Z}\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1}, \tag{34}$$

and $\mathbf{D}_{0Z}$, $\mathbf{D}_{1Z}$, and $\mathbf{D}_{2Z}$ can be obtained as in Eq. (28) from $\mathbf{D}_0$, $\mathbf{D}_1$, and $\mathbf{D}_2$ using the eigenvalues and eigenvectors of $\mathbf{S}$ and $\mathbf{S}^{1/2}\tilde{\mathbf{S}}^{-1}\mathbf{S}^{1/2}$, respectively.

B. The gradient of the matrix U

Collecting the terms, which contain $\partial\mathbf{U}/\partial\lambda$ in Eqs. (21) and (29) or (30), one obtains their total contribution to the NESC energy gradient according to Eq. (35) (superscripts $\tilde{L}$ and G denote contributions from $\tilde{\mathbf{L}}$ and $\mathbf{G}$),

$$\begin{aligned}
&\left(\mathbf{P}\frac{\partial\mathbf{U}}{\partial\lambda} + \mathbf{P}^\dagger\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\right)^{\tilde{L}} + \left(\mathbf{P}^G\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\mathbf{U} + \mathbf{P}^G\mathbf{U}^\dagger\frac{\partial\mathbf{U}}{\partial\lambda}\right)^{G} \\
&= tr\left\{\tilde{\mathbf{P}}\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}[\mathbf{T} - (\mathbf{T}-\mathbf{W})\mathbf{U}] + \tilde{\mathbf{P}}\left[\mathbf{T} - \mathbf{U}^\dagger(\mathbf{T}-\mathbf{W})\right]\frac{\partial\mathbf{U}}{\partial\lambda}\right\} \\
&\quad - \frac{1}{2mc^2}tr\left(\tilde{\mathbf{N}}\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\mathbf{T}\mathbf{U} + \tilde{\mathbf{N}}\mathbf{U}^\dagger\mathbf{T}\frac{\partial\mathbf{U}}{\partial\lambda}\right) \\
&= tr\left(\mathbf{P}_0\frac{\partial\mathbf{U}}{\partial\lambda} + \mathbf{P}_0^\dagger\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\right),
\end{aligned} \tag{35}$$

where $\mathbf{P}^G = \tilde{\mathbf{N}}\mathbf{T}^\dagger$ and $\mathbf{P}_0$ is

$$\mathbf{P}_0 = \tilde{\mathbf{P}}[\mathbf{T} - \mathbf{U}^\dagger(\mathbf{T}-\mathbf{W})] - \frac{1}{2mc^2}\tilde{\mathbf{N}}\mathbf{U}^\dagger\mathbf{T}, \tag{36}$$

and $\tilde{\mathbf{N}}$ can be $\tilde{\mathbf{Z}}$ from Eq. (29) or $\mathbf{D}_3$ from Eq. (34).

For the purpose of determining the derivatives in Eq. (35), we transform Eq. (8) as in Eq. (37),

$$\mathbf{U} = \mathbf{T}^{-1}\left(\left(\mathbf{I} + \frac{1}{2mc^2}\mathbf{U}^\dagger\mathbf{T}\mathbf{U}\mathbf{S}^{-1}\right)^{-1}\tilde{\mathbf{L}} - \mathbf{V}\right), \tag{37}$$

where $\mathbf{I}$ is a unit matrix. Using the identity $(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I} - (\mathbf{I}+\mathbf{A})^{-1}\mathbf{A}$ (Ref. 45), Eq. (37) can be transformed to Eq. (38),

$$\begin{aligned}
\mathbf{U} = \mathbf{T}^{-1}\Bigg(&\mathbf{T}\mathbf{U} + \mathbf{U}^\dagger\mathbf{T} - \mathbf{U}^\dagger(\mathbf{T}-\mathbf{W}) \\
&\mathbf{U} - \frac{1}{2mc^2}\mathbf{U}^\dagger\mathbf{T}\mathbf{U}\mathbf{S}^{-1} \\
&\times\left(\mathbf{I} + \frac{1}{2mc^2}\mathbf{U}^\dagger\mathbf{T}\mathbf{U}\mathbf{S}^{-1}\right)^{-1}\tilde{\mathbf{L}}\Bigg),
\end{aligned} \tag{38}$$

from which with the help of results of Ref. 37, one obtains Eq. (39):

$$\begin{aligned}
\mathbf{U} &= (\mathbf{T}-\mathbf{W})^{-1}\Bigg(\mathbf{T} - \frac{1}{2mc^2}\mathbf{T}\mathbf{U}\mathbf{S}^{-1} \\
&\quad\times\left(\mathbf{I} + \frac{1}{2mc^2}\mathbf{U}^\dagger\mathbf{T}\mathbf{U}\mathbf{S}^{-1}\right)^{-1}\tilde{\mathbf{L}}\Bigg) \\
&= \mathbf{U}^{IORA} - \frac{1}{2mc^2}\mathbf{U}^{IORA}\mathbf{U}\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}} \\
&= \mathbf{U}^{IORA} - \frac{1}{2mc^2}\mathbf{U}^{IORA}\mathbf{U}\mathbf{S}^{-1}(\mathbf{T}\mathbf{U} + \mathbf{V}).
\end{aligned} \tag{39}$$

Differentiating Eq. (39), substituting into Eq. (35), and collecting similar terms, one arrives at Eqs. (40) and (41):

$$\begin{aligned}
\frac{\partial\mathbf{U}}{\partial\lambda} &= \frac{\partial\mathbf{U}^{IORA}}{\partial\lambda}(\mathbf{U}^{IORA})^{-1}\mathbf{U} - \frac{1}{2mc^2}\mathbf{U}^{IORA}\frac{\partial}{\partial\lambda} \\
&\quad\times[\mathbf{U}\mathbf{S}^{-1}(\mathbf{T}\mathbf{U} + \mathbf{V})] = \frac{\partial\mathbf{U}^{IORA}}{\partial\lambda}(\mathbf{U}^{IORA})^{-1}\mathbf{U} \\
&\quad - \frac{1}{2mc^2}\mathbf{U}^{IORA}\Bigg[\frac{\partial\mathbf{U}}{\partial\lambda}\mathbf{S}^{-1}(\mathbf{T}\mathbf{U} + \mathbf{V}) + \mathbf{U}\frac{\partial\mathbf{S}^{-1}}{\partial\lambda} \\
&\quad\times(\mathbf{T}\mathbf{U} + \mathbf{V}) + \mathbf{U}\mathbf{S}^{-1}\left(\frac{\partial\mathbf{T}}{\partial\lambda}\mathbf{U} + \mathbf{T}\frac{\partial\mathbf{U}}{\partial\lambda} + \frac{\partial\mathbf{V}}{\partial\lambda}\right)\Bigg],
\end{aligned} \tag{40}$$

$$\begin{aligned}
&tr\left(\mathbf{P}_0\frac{\partial\mathbf{U}}{\partial\lambda} + \mathbf{P}_0^\dagger\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\right) \\
&= tr(\mathbf{P}_{0S} + \mathbf{P}_{0S}{}^\dagger)\frac{\partial\mathbf{S}}{\partial\lambda} + tr\left(\mathbf{P}_{0T} + \mathbf{P}_{0T}{}^\dagger\right)\frac{\partial\mathbf{T}}{\partial\lambda} \\
&\quad + tr\left(\mathbf{P}_{0V} + \mathbf{P}_{0V}{}^\dagger\right)\frac{\partial\mathbf{V}}{\partial\lambda} + tr\left(\mathbf{P}_{0W} + \mathbf{P}_{0W}{}^\dagger\right)\frac{\partial\mathbf{W}}{\partial\lambda} \\
&\quad + tr\left(\mathbf{P}_1\frac{\partial\mathbf{U}}{\partial\lambda} + \mathbf{P}_1{}^\dagger\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\right)
\end{aligned} \tag{41}$$

$$\begin{aligned}
&= tr\left[\sum_{i=0}^{n}\left(\mathbf{P}_{iS} + \mathbf{P}_{iS}{}^\dagger\right)\right]\frac{\partial\mathbf{S}}{\partial\lambda} + tr\left[\sum_{i=0}^{n}\left(\mathbf{P}_{iT} + \mathbf{P}_{iT}{}^\dagger\right)\right]\frac{\partial\mathbf{T}}{\partial\lambda} \\
&\quad + tr\left[\sum_{i=0}^{n}\left(\mathbf{P}_{iV} + \mathbf{P}_{iV}{}^\dagger\right)\right]\frac{\partial\mathbf{V}}{\partial\lambda} + tr\left[\sum_{i=0}^{n}\left(\mathbf{P}_{iW} + \mathbf{P}_{iW}{}^\dagger\right)\right] \\
&\quad\times\frac{\partial\mathbf{W}}{\partial\lambda} + tr\left(\mathbf{P}_{n+1}\frac{\partial\mathbf{U}}{\partial\lambda} + \mathbf{P}_{n+1}^\dagger\frac{\partial\mathbf{U}^\dagger}{\partial\lambda}\right),
\end{aligned} \tag{42}$$

from which one obtains Eq. (42) by recursion. In Eqs. (41) and (42), the following matrices are used:

$$\mathbf{P}_{i+1} = -\frac{1}{2mc^2}\left(\mathbf{C}_1\mathbf{P}_i\mathbf{U}^{IORA} + \mathbf{P}_i\mathbf{C}_4\right), \tag{43}$$

$$\mathbf{P}_{iS} = \frac{1}{2mc^2}\mathbf{C}_1\mathbf{P}_i\mathbf{C}_2, \tag{44}$$

$$\mathbf{P}_{iT} = \mathbf{C}_5\mathbf{P}_i\mathbf{C}_3 - \frac{1}{2mc^2}\mathbf{U}\mathbf{P}_i\mathbf{C}_2, \tag{45}$$

$$\mathbf{P}_{iV} = -\frac{1}{2mc^2}\mathbf{P}_i\mathbf{C}_2, \tag{46}$$

$$\mathbf{P}_{iW} = \mathbf{U}\mathbf{P}_i\mathbf{C}_3, \tag{47}$$

$$\mathbf{C}_1 = \mathbf{S}^{-1}(\mathbf{T}\mathbf{U} + \mathbf{V}), \tag{48}$$

$$\mathbf{C}_2 = \mathbf{U}^{IORA}\mathbf{U}\mathbf{S}^{-1}, \tag{49}$$

$$\mathbf{C}_3 = \mathbf{U}^{IORA}\mathbf{T}^{-1}, \tag{50}$$

$$\mathbf{C}_4 = \mathbf{C}_2\mathbf{T}, \tag{51}$$

$$\mathbf{C}_5 = -\mathbf{T}^{-1}\mathbf{W}\mathbf{U}. \tag{52}$$

In Eq. (42), $tr(\mathbf{P}_X + \mathbf{P}_X{}^\dagger)\frac{\partial \mathbf{X}}{\partial\lambda}$ $(X = S, T, V, and\ W)$ is not simplified to $2tr(\mathbf{P}_X\frac{\partial \mathbf{X}}{\partial\lambda})$ in order to utilize the triangular storage of the symmetric $(\mathbf{P}_X + \mathbf{P}_X{}^\dagger)$ matrix.

If all eigenvalues of $\tilde{\mathbf{L}}$ are smaller than $2mc^2$, Eq. (42) converges fast and the zeroth order approximation of Eq. (42) (i.e., $n = 0$) is sufficiently accurate. However, if there are some eigenvalues larger than $2mc^2$, Eq. (42) may diverge. In this case, one has to take the zeroth order approximation again and neglect the higher order terms.

C. The gradient of the matrix $\tilde{\mathbf{L}}$

By collecting terms from Eqs. (21) and (42) and substituting them into the first term of Eq. (20), one obtains Eq. (53),

$$\begin{aligned} tr\tilde{\mathbf{P}}\frac{\partial\tilde{\mathbf{L}}}{\partial\lambda} = {} & tr\tilde{\mathbf{P}}''\frac{\partial\mathbf{T}}{\partial\lambda} + tr\tilde{\mathbf{P}}'\frac{\partial\mathbf{W}}{\partial\lambda} + tr\tilde{\mathbf{P}}\frac{\partial\mathbf{V}}{\partial\lambda} + tr(\mathbf{P}_S + \mathbf{P}_S{}^\dagger) \\ & \times\frac{\partial\mathbf{S}}{\partial\lambda} + tr(\mathbf{P}_T + \mathbf{P}_T{}^\dagger)\frac{\partial\mathbf{T}}{\partial\lambda} + tr(\mathbf{P}_V + \mathbf{P}_V{}^\dagger)\frac{\partial\mathbf{V}}{\partial\lambda} \\ & + tr(\mathbf{P}_W + \mathbf{P}_W{}^\dagger)\frac{\partial\mathbf{W}}{\partial\lambda}, \end{aligned} \tag{53}$$

where new matrices

$$\tilde{\mathbf{P}}' = \mathbf{U}\tilde{\mathbf{P}}\mathbf{U}^\dagger \tag{54}$$

and

$$\tilde{\mathbf{P}}'' = \mathbf{U}\tilde{\mathbf{P}} + \tilde{\mathbf{P}}\mathbf{U}^\dagger - \mathbf{U}\tilde{\mathbf{P}}\mathbf{U}^\dagger \tag{55}$$

are introduced. The auxiliary matrices $\mathbf{P}_X$ with X = S, T, V, and W are obtained by appropriate summation of the matrices in Eq. (42). They can be easily calculated at the end of the iterative solution of the NESC equations and then contracted with the one-electron integral derivatives in Eq. (53), which are directly available in non-relativistic quantum chemical codes.

D. Contracted basis sets

When deriving the above formalism, it was tacitly assumed that a basis set of primitive atom-centered basis functions is employed. With the use of contracted basis sets as defined in Eq. (56),

$$\boldsymbol{\chi}^{(c)} = \boldsymbol{\chi}^{(p)}\mathbf{R}, \tag{56}$$

the NESC equation (1) can be solved in the basis of the primitive functions $\boldsymbol{\chi}^{(p)}$ and then the NESC one-electron Hamiltonian equation (12) can be transformed to the contracted basis set $\boldsymbol{\chi}^{(c)}$ via Eq. (57),

$$\mathbf{H}^{(c)}_{1-e} = \mathbf{R}^\dagger\mathbf{H}^{(p)}_{1-e}\mathbf{R}, \tag{57}$$

where the Hamiltonian $\mathbf{H}^{(p)}_{1-e}$ is calculated over the primitive basis functions and $\mathbf{R}$ is a rectangular $(N \times M)$ matrix of contraction coefficients (N primitives, M contracted basis functions).

With the use of a contracted basis set, the first term in Eq. (17) can be evaluated according to Eq. (58):

$$\begin{aligned} \frac{\partial E_{1-e}}{\partial\lambda} &= tr\mathbf{P}^{(c)}\left(\frac{\partial\mathbf{H}^{(c)}_{1-e}}{\partial\lambda}\right) \\ &= tr\mathbf{P}^{(c)}\left(\mathbf{R}^\dagger\frac{\partial\mathbf{H}^{(p)}_{1-e}}{\partial\lambda}\mathbf{R}\right) \\ &= tr\left(\mathbf{R}\mathbf{P}^{(c)}\mathbf{R}^\dagger\right)\left(\frac{\partial\mathbf{H}^{(p)}_{1-e}}{\partial\lambda}\right), \end{aligned} \tag{58}$$

where the derivatives of the NESC one-electron Hamiltonian $\mathbf{H}^{(p)}_{1-e}$ are calculated as described above. Thus, in Eq. (20), it is only necessary to replace the density matrix $\mathbf{P}$ by $\mathbf{R}\mathbf{P}^{(c)}\mathbf{R}^\dagger$ when using contracted basis sets.

E. Arguments for neglecting the derivative of U

Any simplification of the NESC gradient implies a careful evaluation of the magnitude of the derivatives collected in $\partial\mathbf{U}\,/\,\partial\lambda$. For this purpose, we start with $\mathbf{U}$ in the form of Eq. (39). Applying recursively the matrix identity $(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I} - (\mathbf{I}+\mathbf{A})^{-1}\mathbf{A}$ (Ref. 45) to the expression (7), one obtains the following expansion for the $\mathbf{U}^{IORA}$ operator:

$$\mathbf{U}^{IORA} = \mathbf{I} + \mathbf{T}^{-1}\mathbf{W} + \mathbf{T}^{-1}\mathbf{W}\mathbf{T}^{-1}\mathbf{W} + \ldots. \tag{59}$$

The terms on the right hand side of Eq. (59) are of the order c^{-2}, c^{-4}, etc. Combining Eqs. (39) and (59) leads to Eq. (60):

$$\mathbf{U} = \mathbf{I} + \left(\mathbf{T}^{-1}\mathbf{W} - \frac{1}{2mc^2}\mathbf{U}^{IORA}\mathbf{U}\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{L}}\right) + \mathbf{T}^{-1}\mathbf{W}\mathbf{T}^{-1}\mathbf{W} + \ldots, \tag{60}$$

where the two terms in parentheses have the same order of magnitude. Because the matrix $\mathbf{W}$ is negative definite, it may be expected that the two terms in parentheses in Eq. (60) should to some degree compensate each other, at least for the bound (negative energy) eigenvalues of the operator $\tilde{\mathbf{L}}$. Furthermore, the terms on the right hand side of Eq. (60) become significant for the very tight basis functions, which are important for the total energy. Neglecting them in the NESC equations (that is setting $\mathbf{U} = \mathbf{I}$) leads to variational collapse of the method.[46] Since in a typical relativistic calculation on molecules with heavy elements, steep basis functions are always included, the derivative of $\mathbf{U}$ with respect to the nuclear coordinates cannot *a priori* be excluded although its contribution should be small in many cases. This however changes for other first order properties, which depend more strongly on the density closer to the nucleus and, therefore,

TABLE I. Specification of relativistic basis sets used in this work.[a]

Element	Description	Reference
A. CCSD and CCSD(T) calculations		
Hg	(22s19p12d9f)/[15s13p8d5f].	14
H,C,N,O,F,Cl,Br	Dyall's Dirac-contracted cc-pVTZ(fi/sf/fw) with diffuse functions taken from Dunning's aug-cc-pVTZ.	52
I	All-electron relativistic def2-TZVPP basis; the core functions were re-contracted at the SA-CASSCF level using the NESC Hamiltonian.	49 and 50
B. DFT(B3LYP) calculations of BDT-Hg and BDET-Hg		
Hg	All-electron relativistic SARC basis; the core functions were re-contracted at the Hartree-Fock level using the NESC Hamiltonian.	49 and 50
H,C,N,O	Re-contracted 6-311G basis at the SA-CASSCF level using the NESC Hamiltonian.	52
S,Cl	Dyall's Dirac-contracted cc-pVTZ(fi/sf/fw) basis with diffuse functions taken from Dunning's aug-cc-pVTZ.	52
C. DFT(PBE0) calculations of TlX and Hg_2		
Tl,Hg	All-electron relativistic SARC basis; the core functions were re-contracted at the Hartree-Fock level using the NESC Hamiltonian. For Tl, the basis functions were augmented by 3g2h1i functions.	49, 50, and 51
F,Cl,Br	All-electron def2-QZVPP basis; the core functions were re-contracted at the SA-CASSCF level using the NESC Hamiltonian.	52
I	All-electron relativistic def2-TZVPP basis; the core functions were re-contracted at the SA-CASSCF level using the NESC Hamiltonian.	49 and 50

[a]All calculations, if not noted otherwise, with a finite nucleus model.

we have included $\partial\mathbf{U}/\partial\lambda$ into the standard version of our NESC gradient program.

On the other hand, contributions of $\partial\mathbf{U}/\partial\lambda$ to $\partial\tilde{\mathbf{S}}/\partial\lambda$ (see the third term in Eq. (26)) can be neglected even in the case of steep functions. The term containing $\partial\mathbf{U}/\partial\lambda$ is of the order c^{-4} and accordingly it is significantly smaller than the first two terms in Eq. (26). Neglecting the contribution from $\partial\mathbf{U}/\partial\lambda$ does not lead to any substantial loss of accuracy.

F. Deriving the NESC gradient for a finite nucleus model

When calculating the derivatives $\partial\mathbf{V}/\partial\lambda$ and $\partial\mathbf{W}/\partial\lambda$, both point charge and finite nucleus model with Gaussian charge distribution may be used. In this work, we determined the derivatives of $\mathbf{W}$ for the finite nucleus model, whereas in the case of $\mathbf{V}$ the point charge model was assumed because the differences between the two models are negligible in the latter case. When calculating molecular geometries, the effect of the nuclear model on the gradient is negligible, in general. This is a result of the fact that the changes of $\mathbf{V}$ and $\mathbf{W}$ in the bond regions of a molecule are small. However, for other properties, especially those that depend on the density close to the nucleus, the nuclear model used in both $\partial\mathbf{W}/\partial\lambda$ and $\partial\mathbf{V}/\partial\lambda$ may have a significant impact on the NESC derivatives.

IV. COMPUTATIONAL TECHNIQUES

The algorithms described above have been programmed within the COLOGNE2010 program package.[47] Each part of the analytical energy gradient was checked against the corresponding numeric approximation. For this purpose, very accurate calculations were required that used an ultrafine grid[48] for the density functional theory (DFT) calculations, a SCF convergence criterion of 10^{-8} and a geometry optimization convergence criteria of 10^{-7} with regard to the mean force, the mean displacement, and the absolutely largest force or displacement. For comparison purposes, bond lengths determined by either the analytical NESC energy gradient or its numerical equivalent were accurate up to four decimal places.

NESC calculations, which demonstrate the application potential of the analytical energy gradient, were carried out with a variety of basis sets[14,49–52] (compare with Table I), which in some cases had to be re-contracted for the core functions as indicated in Table I. From our previous experience with these and similar basis sets,[14,15,53,54] the basis set superposition error (BSSE) amounts to less than ~1 kcal/mol (see also BSSE corrections in Tables II–III). We used DFT to check and compare features of the NESC gradient (see Table II). In this connection, we applied the PBE0 exchange-correlation (XC) functional[55,56] to investigate the thallium halogenides **1**–**4** of Table II.

NESC/CCSD (coupled cluster with all single and double excitation) (Ref. 57) geometry optimizations followed by NESC/CCSD(T) (CCSD with a perturbative inclusion of all triple excitations; Ref. 58) single point calculations were performed in the case of the mercury molecules **5**–**13** of Table III. In addition, we calculated the geometries of the organic mercury molecules **14** and **15** (see Fig. 1) at the NESC/DFT level of theory utilizing the B3LYP hybrid functional[59–62] to demonstrate the feasibility of large NESC calculations. The largest calculations carried out in this work were for $HgCF_3$ at the NESC/CCSD(T) level of theory with 434 primitive basis functions and 313 contracted basis functions and at the NESC/B3LYP level of theory for molecule **14** (primitives: 786; contracted: 417 basis functions), where it has to be noted that the NESC equations are solved in the basis of the primitives ("first-diagonalize-then-contract").

All calculations were carried out with a finite nucleus model possessing a Gaussian charge distribution.[63,64] Furthermore, the renormalization of the one-electron Hamiltonian

TABLE II. Bond lengths r_e, bond dissociation energies D_e (D_0), and excitation energies T_e of the $X^1\Sigma^+$ ground state of thallium monohalogenides.[a]

Molecule	Method	Bond length r_e	D_e	D_0	$T_e(^3\Pi - X^1\Sigma^+)$	Reference
TlF (**1**)	NESC/PBE0					
	no dU	2.09591	117.0	102.4	4.39	This work
	dU	2.09592				
	num.	2.0959				
	dU^{IORA}	oscillating				
	Expt.	2.0844		105.4	$A0^+$: 4.36; B1: 4.57	68 and 69
	RECP/CCSD(T)	2.078		106.8		70
	RECP/CASPT3(CI)	2.087		105.4		68
TlCl (**2**)	NESC/PBE0					
	no dU	2.50442	101.4	86.4	3.94	This work
	dU	2.50445				
	num.	2.5045				
	dU^{IORA}	oscillating				
	Expt .	2.4848		88.1	$A0^+$: 3.85	68 and 69
	RECP/CCSD(T)	2.486		87.2		70
	RECP/CASPT3(CI)	2.487		87.9		68
TlBr (**3**)	NESC/PBE0					
	no dU	2.64174	93.1	76.1	3.69	This work
	dU	2.64178				
	num.	2.6418				
	dU^{IORA}	oscillating				
	Expt .	2.618 2		78.8	$A0^+$: 3.62	68 and 69
	RECP/CCSD(T)	2.618		76.5		70
	RECP/CASPT3(CI)	2.621		75.4		68
TlI (**4**)	NESC/PBE0					
	no dU	2.84166	83.1	61.2	3.36	This work
	dU	2.84166				
	num.	2.8417				
	dU^{IORA}	oscillating				
	Expt .	2.8137		63.7	$A0^+$: 2.99	68 and 69
	RECP/CCSD(T)	2.813		62.5		70
	RECP/CASPT3(CI)	2.825		63.0		68

[a]D_0 values were obtained by correcting D_e values for zero-point energy (ZPE) and spin-orbit coupling (SOC) taken from Ref. 68. Bond lengths in Å, D values in kcal/mol, excitation energies T_e in eV. The following abbreviations are used: no dU: NESC gradient calculated without including $\partial\mathbf{U}/\partial\lambda$; dU: NESC gradient calculated including $\partial\mathbf{U}/\partial\lambda$; num.: numerical NESC gradient used; dU^{IORA}: $\partial\mathbf{U}/\partial\lambda$ is approximated by $\partial\mathbf{U}^{IORA}/\partial\lambda$; RECP: relativistic effective core potential; CASPT3(CI): configuration interaction based complete active space third order perturbation theory. - Calculated BSSE corrections (not included) are −0.12 (F), −0.09 (Cl), −0.22 (Br), and −0.83 kcal/mol (I).

according to Eq. (15) ("picture change" correction[41]) was applied throughout this work. For the scalar relativistic NESC approach, a velocity of light $c = 137.035999070(98)$ (Ref. 65) was used throughout the article.

V. RESULTS AND DISCUSSIONS

Table II contains NESC/PBE0 bond lengths, D_e (D_0) dissociation energies, and excitation energies T_e of thallium halides in their ground state. In Table III, exact quasi-relativistic 1c-NESC geometries obtained at the NESC/CCSD and bond dissociation energies (D_e and D_0) calculated at the NESC/CCSD(T) level of theory are listed and compared with results of other calculations as well as the experimental data.[68–70] Also summarized in Table III are some NESC/DFT results calculated for the two organic mercury compounds **14** and **15** shown in Fig. 1.

Differences between bond lengths TlX obtained with the analytical NESC energy gradient and those calculated with a numeric gradient are smaller than 10^{-4} Å (Table II) and by this within the errors of a finite differences evaluation of first derivatives. Calculations reveal also that an approximation of $\partial\mathbf{U}/\partial\lambda$ by $\partial\mathbf{U}^{IORA}/\partial\lambda$ cannot be recommended because in most cases oscillations rather than convergence was observed when solving the NESC equations. If the analytical form of $\partial\mathbf{U}/\partial\lambda$ is not available, it is preferable to neglect this term rather than to approximate it. In the case of the TlX molecules, deviations in the bond lengths are 10^{-5} Å if $\partial\mathbf{U}/\partial\lambda$ is neglected. These deviations, however, increase already to 10^{-3} Å if distances between heavy atoms are investigated with normal basis sets. An example is the Hg,Hg distance in the Hg_2 van der Waals complex (3.587 vs 3.588 Å, NESC/PBE0). The differences can take values of even 10^{-2} Å if very steep basis functions with exponents larger than 10^9 are involved. Since the calculation of $\partial\mathbf{U}/\partial\lambda$ does not involve a major computational effort, all calculations are carried out with the correct calculation of $\partial\mathbf{U}/\partial\lambda$.

Calculated TlX bond lengths reasonably agree with the known experimental bond lengths where deviations increase from 0.015 (X = F) to 0.027 Å (X = I), which may indicate the increasing influence of spin orbit coupling (SOC) on the bond length, which was not considered in this work. Apart

TABLE III. NESC/CCSD, NESC/CCSD(T), and NESC/DFT descriptions of mercury molecules.[a] – Calculated BSSE corrections (not included) of HgF and HgCl are −0.40 and −0.53 kcal/mol, respectively.

Molecule	Symmetry	state	method	Geometry parameters	D_e (D_0)	Reference
HgF (**5**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	2.024	33.0 (32.3)	This work
			IORA/QCISD	2.025		54
			NESC(d)/CCSD(T)//B3LYP	2.080	31.0	14
			Expt.		32.9	69
HgCl (**6**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	2.402	23.8 (23.4)	This work
			SOC/RECP/CCSD(T)	2.354	22.9	71
			NESC(d)/CCSD(T)//B3LYP	2.460	21.6	14
			Expt.	2.395, 2.42	23.4, 24.6	72, 73, and 74
HgBr (**7**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	2.546	20.0 (17.5)	This work
			SOC/RECP/CCSD(T)	2.498	16.3	71
			NESC(d)/CCSD(T)//B3LYP	2.672	14.7	14
			Expt.	2.62	17.2, 18.4	75, 76, and 77
HgI (**8**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	2.709	12.9 (7.6)	This work
			SOC/RECP/CCSD(T)	2.708	8.6	71
			NESC(d)/CCSD(T)//B3LYP	2.820	11.2	14
			Expt.	2.81	7.8, 8.1, 8.9	73, 74, and 78
HgCN (**9**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	Hg-C: 2.118, C-N: 1.161	36.1	This work
			IORA/QCISD	Hg-C: 2.114, C-N: 1.179		54
			NESC(d)/CCSD(T)//B3LYP	Hg-C: 2.159	32.6	14
HgNC (**10**)	$C_{\infty v}$	$^2\Sigma^+$	NESC/CCSD(T)//NESC/CCSD	Hg-N: 2.077, N-C: 1.176	22.4	This work
$HgCH_3$ (**11**)	C_{3v}	$^2A1$	NESC/CCSD(T)//NESC/CCSD	Hg-C: 2.344, H-C: 1.084, Hg-C-H: 104.3	3.2	This work
			NESC(d)/CCSD(T)//B3LYP	Hg-C: 2.455	3.0	14
$HgOCH_3$ (**12**)	C_s	$^2A'$	NESC/CCSD(T)//NESC/CCSD	Hg-O: 3.116, C-O: 1.379, C-H: 1.094, 1.094, 1.101, Hg-O-C: 107.4	2.9	This work
$HgCF_3$ (**13**)	C_{3v}	$^2A1$	NESC/CCSD(T)//NESC/CCSD	Hg-C: 2.441, F-C: 1.327, Hg-C-F: 110.0	9.6	This work
			NESC(d)/CCSD(T)//B3LYP	Hg-C: 2.617	2.7	14
BDT-Hg (**14**)	C_2	^1A	NESC/B3LYP	Hg-S: 2.356, S-Hg-S: 175.9		This work
			Expt.	Hg-S: 2.42, S-Hg-S: ∼180		66
BDET-Hg (**15**)	C_1	^1A	NESC/B3LYP	Hg-S: 2.378, 2.387 S-Hg-S: 163.2		This work
			Expt.	Hg-S: 2.327, 2.340 S-Hg-S: 166.8		67

[a]NESC/CCSD(T)//NESC/CCSD denotes NESC/CCSD(T) energies calculated at NESC/CCSD geometries; NESC(d): NESC calculations based on an iterative calculation of **U** with damping (d); QCISD: quadratic configuration interaction with single and double excitations. For HgX (X = F, Cl, Br, I), D_0 values corrected by ZPE and SOC values are given in parentheses (see footnote to Table II). D values in kcal/mol, bond lengths in Å, angles in deg. In the calculations of $HgCH_3$, $HgOCH_3$, and $HgCF_3$, the fourteen 4f-electrons of Hg were frozen. NESC(d) calculations were carried out with the Filatov-Dyall version of NESC based on damping (d) of the **U** matrix.[13] For the structure of molecules **16** and **17**, compare with Figure 1.

from this, the limitations of the XC-functional are the major cause for the differences. Calculated and measured D_0 values differ on the average by −2.5 kcal/mol. Since the computed values were combined with zero-point energy (ZPE) and SOC corrections, there is no increase in the deviations with increasing atomic number of X. The largest deviation is found for X = F (−3.0 kcal/mol, Table II). It has to be stressed that the agreement between theory and experiment is excellent in view of the large experimental D_0 values of 64 (X = I) to 102 kcal/mol (X = F) and errors of just 3-4 % of the experimental values.

Experimental singlet-triplet excitation energies of TlX are well reproduced for X = F, Cl, and Br for which deviations are smaller than 0.1 eV (Table II). Only in the case of TlI, a large deviation of 0.37 eV is calculated at the NESC/PBE0 level. Contrary to the D_0 values, all calculated excitation energies T_e are larger than the corresponding experimental values.

NESC/CCSD geometries of mercury containing molecules are partly in agreement with other published geometries,[14,54,71] such as those obtained by IORA/QCISD (HgF) or SOC/RECP/CCSD(T) (HgI), partly they are closer to the experiment as in the case of HgCl or HgBr. It is difficult to make a judgment on the accuracy of the NESC calculations, because the influence of SOC on the bond length was not considered. However, in all cases where a comparison is possible, the difference of NESC HgX bond lengths from experimental values is smaller than 0.07 Å with the exception of HgI (difference of 0.1 Å). Agreement between NESC/CCSD(T) and experimental D_0 dissociation energies for mercury halides[72–78] is excellent in view of a mean deviation of just 0.3 kcal/mol. This agreement is

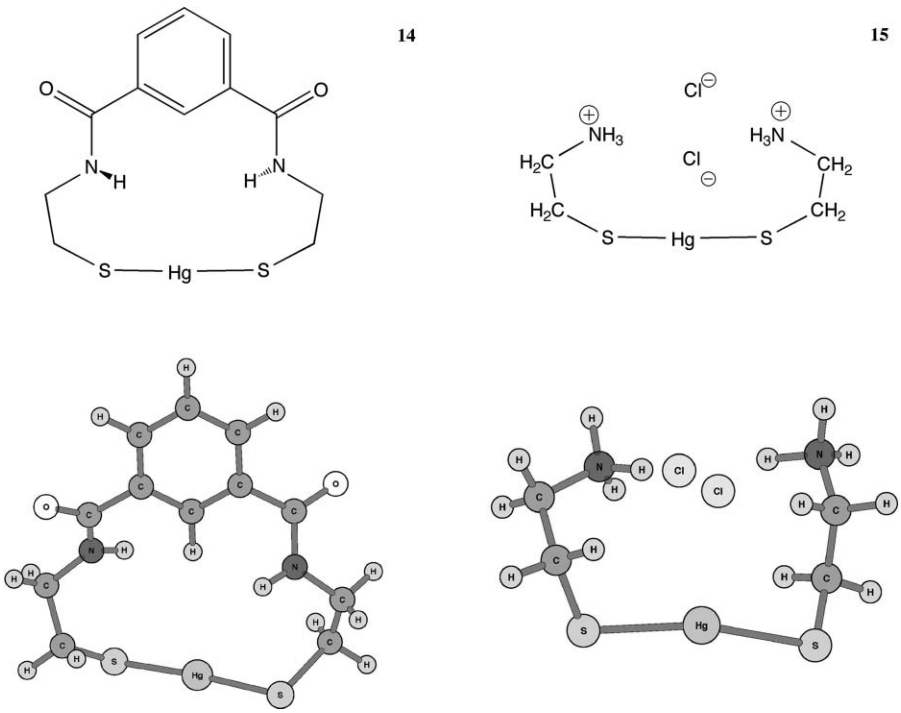

FIG. 1. Structure of organic mercury molecules **14** and **15**. The NESC/B3LYP geometries are given in the lower half of the figure.

obtained when selecting that experimental D_0 value which is closest to the NESC/CCSD(T) result, i.e., NESC/CCSD(T) provides reliable values so that the most accurate experimental D_0 value can be identified. It is interesting to note that the renormalization of the one-electron Hamiltonian according to Eq. (15), known in the literature as "picture change,"[41] leads to an increase of calculated D_e values for HgX (X= F, Cl, Br, I) by 0.2, 0.2, 0.2, 0.1 kcal/mol, respectively. Hence, if high accuracy is needed, this contribution cannot be excluded.

We calculated also the two organo-mercury compounds **14** and **15** shown in Fig. 1. Compound **14** is obtained from the reaction of 1,3-benzenediamidoethanethiol ($BDTH_2$) with $HgCl_2$ in water leading to BDT-Hg **14**,[66] which is known to provide an excellent means to remove mercury permanently from water. A similar reaction of bisaminoethanthiol $BDETH_2$ with $HgCl_2$ yields BDET-Hg **15**.[67] NESC/B3LYP geometry optimizations of **14** and **15** took less than 1 h (using 8 processors) and provided data in close agreement with the x-ray analysis of these compounds (Table III and Fig. 1). In forthcoming work, we will investigate organo-mercury compounds being of relevance for the removal of mercury from the environment.

VI. CONCLUSIONS

In this work, we have described the derivation, implementation, and application of the analytical NESC energy gradient.

(1) We have outlined an effective algorithm to calculate the NESC energy also in the presence of basis functions with very large exponent ("steep basis functions"). Such an algorithm is the prerequisite for exact quasi-relativistic NESC calculations in combination with HF, DFT, or correlation corrected *ab initio* calculations, such as CI, MPn perturbation theory, coupled cluster theory, complete active space self-consistent field (CASSCF), or any other non-relativistic method.

(2) We have developed the analytical NESC energy gradient in its complete form, which implies the derivation of the analytic form of the derivative $\partial\mathbf{U}/\partial\lambda$, where **U** is the matrix representation of the operator $\hat{U}$ for the elimination of the small component. The derivative of **U** is rather complex and contributes to different parts of the NESC energy gradient so that its efficient calculation was the key of the current work.

(3) For basis sets with very steep basis functions it is advisable to calculate $\partial\mathbf{U}/\partial\lambda$, whereas for other basis sets $\partial\mathbf{U}/\partial\lambda$ may be neglected. In this work, we have discussed the reasons for the different magnitudes of $\partial\mathbf{U}/\partial\lambda$.

(4) Corrections for the picture change were programmed and tested for the NESC energy as well as the NESC energy gradient. There is a small contribution of 0.2 (F) up to 0.1 kcal/mol (I) in calculated dissociation energies D_e of mercury halides. For the calculated geometries, the changes due to picture change corrections are of the order of 10^{-4} Å or smaller.

(5) The analytical energy gradient was also developed for a finite nucleus model with a Gaussian distribution of positive charge. The influence of the nuclear model on the NESC energy is significant, whereas it is negligible for geometry calculations.

(6) The economic implementation of the analytic NESC energy gradient and its general applicability is emphasized by presenting NESC/B3LYP geometry optimizations using up to 800 primitive basis functions ("first-diagonalize-then-contract" strategy) for organic mercury compounds.

(7) NESC/CCSD(T) bond dissociation enthalpies obtained at NESC/CCSD geometries and corrected for spin orbit coupling and zero-point energies agree well with experimental D_0 values and offer the possibility of identifying the most reliable experimental value if several are available.
(8) NESC/DFT has been used to determine bond lengths and D_0 dissociation energies at 298 K for thallium halides with the result that the former differ by 0.03 Å or less whereas the latter underestimate experimental values on the average by 2.5 kcal/mol.

Summarizing, we have proven that NESC and NESC gradient calculations can routinely be carried out and that, with the development of analytical first order energy derivatives, the applicability range and the usefulness of NESC has been substantially increased. In separate work, we will extend the applicability of NESC to second order response properties.

ACKNOWLEDGMENTS

This work was financially supported by the National Science Foundation, Grant No. CHE 071893. We thank SMU for providing computational resources.

[1] P. A. M. Dirac, Proc. R. Soc. London A **117**, 610 (1928).
[2] P. A. M. Dirac, Proc. R. Soc. London A **123**, 714 (1929).
[3] K. G. Dyall, J. Chem. Phys. **106**, 9618 (1997).
[4] K. G. Dyall, J. Comput. Chem. **23**, 786 (2002).
[5] A. Wolf, M. Reiher, and B. A. Hess, J. Chem. Phys. **117**, 9215 (2002).
[6] A. Wolf, M. Reiher, and B. Hess, J. Chem. Phys. **120**, 8624 (2004).
[7] M. Reiher and A. Wolf, J. Chem. Phys. **121**, 2037 (2004).
[8] M. Iliaš and T. Saue, J. Chem. Phys. **126**, 064102 (2007).
[9] M. Barysz, L. Mentel, and J. Leszczynski, J. Chem. Phys. **130**, 164114 (2009).
[10] W. Kutzelnigg and W. Liu, J. Chem. Phys. **123**, 241102 (2005).
[11] W. Kutzelnigg and W. Liu, Mol. Phys. **104**, 2225 (2006).
[12] W. Liu and W. Kutzelnigg, J. Chem. Phys. **126**, 114107 (2007).
[13] M. Filatov and K. G. Dyall, Theor. Chem. Acc. **117**, 333 (2007).
[14] D. Cremer, E. Kraka, and M. Filatov, ChemPhysChem **9**, 2510 (2008).
[15] E. Kraka, M. Filatov, and D. Cremer, Croat. Chem. Acta **82**, 233 (2009).
[16] E. Kraka and D. Cremer, Int. J. Mol. Sci. **9**, 926 (2008).
[17] M. Filatov and D. Danovich, J. Phys. Chem. A **112**, 12995 (2008).
[18] M. Filatov, J. Chem. Phys. **127**, 084101 (2007).
[19] R. Kurian and M. Filatov, J. Chem. Theory Comput. **4**, 278 (2008).
[20] R. Kurian and M. Filatov, J. Chem. Phys. **130**, 124121 (2009).
[21] R. Kurian and M. Filatov, Phys. Chem. Chem. Phys. **12**, 2758 (2010).
[22] M. Douglas and N. M. Kroll, Ann. Phys. (N.Y.) **82**, 89 (1974).
[23] B. A. Hess, Phys. Rev. A **32**, 756 (1985).
[24] B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[25] G. Jansen and B. A. Hess, Phys. Rev. A **39**, 6061 (1989).
[26] R. Samzow, B. A. Hess, and G. Jansen, J. Chem. Phys. **96**, 1227 (1992).
[27] T. Nakajima and K. Hirao, J. Chem. Phys. **113**, 7786 (2000).
[28] C. van Wüllen, J. Chem. Phys. **120**, 7307 (2004).
[29] M. Reiher and A. Wolf, J. Chem. Phys. **121**, 10945 (2004).
[30] M. Barysz, A. J. Sadlej, and J. G. Snijders, Int. J. Quantum Chem. **65**, 225 (1997).
[31] M. Barysz and A. J. Sadlej, J. Chem. Phys. **116**, 2696 (2002).
[32] M. Barysz and A. Sadlej, J. Chem. Phys. **116**, 2696 (2003).
[33] D. Kedziera and M. Barysz, Chem. Phys. Lett. **393**, 521 (2004).
[34] D. Kedziera, J. Chem. Phys. **123**, 074109 (2005).
[35] D. Kedziera and M. Barysz, Chem. Phys. Lett. **446**, 176 (2007).
[36] M. Iliaš, H. J. A. Jensen, B. O. Roos, and M. Urban, Chem. Phys. Lett. **408**, 210 (2005).
[37] M. Filatov, J. Chem. Phys. **125**, 107101 (2006).
[38] W. Kutzelnigg and W. Liu, J. Chem. Phys. **125**, 107102 (2006).
[39] M. Filatov and D. Cremer, J. Chem. Phys. **122**, 064104 (2005).
[40] J. Jansson and A. Logg, ACM Trans. Math. Softw. **35**, 17 (2008).
[41] W. Liu and D. Peng, J. Chem. Phys. **131**, 031104 (2009).
[42] P. Pulay, Mol. Phys. **17**, 197 (1969).
[43] M. Filatov and D. Cremer, J. Chem. Phys. **118**, 6741 (2003).
[44] V. A. Nasluzov and N. Rösch, Chem. Phys. **210**, 413 (1996).
[45] H. V. Henderson and S. R. Searle, SIAM Rev. **23**, 53 (1981).
[46] K. G. Dyall, J. Chem. Phys. **109**, 4201 (1998).
[47] E. Kraka, J. Gräfenstein, M. Filatov, *et al.*, COLOGNE2010 (2010).
[48] V. I. Lebedev and L. Skorokhodov, Dokl. Math. **45**, 587 (1992).
[49] ORCA, an *ab initio*, Density Functional, and Semiempirical Program Package, version 2.8, F. Neese, F. Wennmohs, *et al.*, Universität Bonn, Bonn, Germany, 2010, also see http://www.thch.uni-bonn.de/tc/orca/.
[50] D. A. Pantazis, X.-Y. Chen, C. R. Landis, and F. Neese, J. Chem. Theory Comput. **4**, 908 (2008).
[51] K. A. Peterson and K. E. Yousaf, J. Chem. Phys. **133**, 174116 (2010).
[52] See https://bse.pnl.gov/bse/portal for EMSL Basis Set Exchange.
[53] M. Filatov and D. Cremer, ChemPhysChem. **5**, 1547 (2004).
[54] M. Filatov and D. Cremer, J. Chem. Phys. **121**, 5618 (2004).
[55] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[56] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1998).
[57] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[58] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[59] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[60] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[61] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[62] P. J. Stevens, F. J. Devlin, C. F. Chablowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[63] L. Visscher and K. G. Dyall, At. Data Nucl. Data Tables **67**, 207 (1997).
[64] K. G. Dyall and K. Fægri, *Introduction to Relativistic Quantum Chemistry* (Oxford University Press, Oxford, 2007).
[65] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, Phys. Rev. Lett. **97**, 030802 (2006).
[66] L. Y. Blue, P. Jana, and D. A. Atwood, Fuel **89**, 1326 (2010).
[67] C.-H. Kim, S. Parkin, Y. S. Lee, and D. A. Atwood, Polyhedron **21**, 225 (2002).
[68] W. Zou and W. Liu, J. Comput. Chem. **30**, 524 (2009).
[69] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure., IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
[70] B. Metz, M. Schweizer, H. Stoll, M. Dolg, and W. Liu, Theor. Chem. Acc. **104**, 22 (2000).
[71] B. C. Shepler, N. B. Balabanov, and K. A. Peterson, J. Phys. Chem. A **109**, 10363 (2005).
[72] A. K. Rai, S. B. Rai, and D. K. Rai, J. Phys. B **15**, 3239 (1982).
[73] J. Tellinghuisen, P. C. Tellinghuisen, S. A. Davies, P. Berwanger, and K. S. Viswanathan, Appl. Phys. Lett. **41**, 789 (1982).
[74] B. E. Wilcomb and R. B. Bernstein, J. Mol. Spectrosc. **62**, 442 (1976).
[75] G. Ullas, S. B. Rai, and D. K. Rai, J. Phys. B **25**, 4497 (1992).
[76] J. Tellinghuisen and J. G. Ashmore, Appl. Phys. Lett. **40**, 867 (1982).
[77] J. Tellinghuisen and J. G. Ashmore, Chem. Phys. Lett. **102**, 10 (1983).
[78] C. Salter, P. C. Tellinghuisen, J. G. Ashmore, and J. Tellinghuisen, J. Mol. Spectrosc. **120**, 334 (1986).