Cost-effective treatment of scalar relativistic effects for multireference systems: a CASSCF implementation based on the spin-free Dirac-Coulomb Hamiltonian

Filippo Lipparini* and Jürgen Gauss

Institut für Physikalische Chemie, Universität Mainz, Duesbergweg 10-14, D-55128 Mainz, Germany

E-mail: flippari@uni-mainz.de

Abstract

We present an implementation of the complete active space–self-consistent field (CASSCF) method specifically designed to be used in four-component scalar relativistic calculations based on the spin-free Dirac-Coulomb (SFDC) Hamiltonian. Our implementation takes full advantage of the properties of the SFDC Hamiltonian that allow us to use real algebra and to exploit point-group and spin symmetry to their full extent while including in a rigorous way scalar relativistic effects in the treatment. The SFDC-CASSCF treatment is more expensive than its non-relativistic counterpart only in the orbital optimization step, while exhibiting the same computational cost for the rate-determining full configuration interaction part. The numerical aspects are discussed and the capabilities of the SFDC-CASSCF methodology are demonstrated through a pilot application.

1 Introduction

Relativistic effects play an important role in atomic and molecular systems.[1–9] The inclusion of relativity in a quantum chemical description is therefore an important aspect of modern computational chemistry, especially when the target of the investigation is a system that includes heavy elements.

The most rigorous way of achieving the inclusion of relativistic effects in a quantum chemical description is to recast the electronic structure methods in a relativistic fashion by formulating them for the full, four-component Dirac-Coulomb (DC) Hamiltonian, possibly with the inclusion of correction terms to the electron-electron interaction such as the Breit contribution.[10] In the last decades, many methods have been implemented for the DC Hamiltonian, including those based on Hartree-Fock[11–14] (HF), Kohn-Sham density functional,[15–18] and multiconfigurational self-consistent field[19–21] (MCSCF) theories. Furthermore, many post-Hartree-Fock methods such as second-order Møller-Plesset perturbation theory[22,23] (MP2), configuration interaction[24–26] (CI), and coupled cluster[27–29] (CC) theory have been implemented for the DC Hamiltonian.

Unfortunately, the use of the DC Hamiltonian introduces several computational overheads. This is mainly due to three concurrent factors. The DC Hamiltonian introduces a four-component spinor and requires one to take into account the small component wave function, it does not admit a real matrix representation, and it does not retain spin and point-group spatial symmetry as in the non-relativistic case. Because of these characteristics, in order to perform a DC quantum chemical calculation, complex algebra must be used. Furthermore, the number of integrals that needs to be computed is highly increased due to the presence of the small component wave function. The small component integrals are evaluated as integral derivatives of the large component ones, which makes them more expensive. Also, because of the SO interaction, many more intermediate quantities need to be assembled in order to compute the two electron integrals. The orthogonality of α and β spin functions can no longer be exploited because of the lack of spin symmetry, which together

with the fact that the wave function can not be chosen to be real-valued and with the lowering of permutational symmetry of the MO integrals, causes a large increase in the number of independent MO integrals. A detailed discussion of the various factors that increase the computational cost of a DC relativistic calculation can be found in ref. 30.

For all these reasons, many alternative approaches to treat relativistic effects have been suggested in the literature.[31–40] Nevertheless, retaining a rigorous, four-component treatment is desirable. A promising way to achieve a cost-effective, but rigorous, treatment of relativistic effects is based on the observation that spin-orbit (SO) interactions are much smaller in magnitude than scalar relativistic effects.[1] As the most severe effects on the computational cost of a four-component calculation stem from the presence of the spin-orbit interaction, a viable strategy would be to decouple it from the spin-free (SF) relativistic effects by decomposing the DC Hamiltonian into a sum of a purely SF and a purely SO term and then discarding the latter term. Such a strategy can be based on the spin separation of the DC Hamiltonian proposed by Dyall[41] and leads to a spin-free Dirac-Coulomb (SFDC) Hamiltonian. The spin separation approach has been explored both at the self-consistent field (SCF) and at correlated levels.[30,42–45] In this contribution, we present an implementation of the complete active space–self-consistent field[46,47] (CASSCF) method specifically tailored for the SFDC Hamiltonian. A relativistic, multireference quantum chemical method is, for example, needed to provide a qualitatively correct description of heavy transition metal complexes, where relativistic effects play an important role and which are characterized by a complex electronic structure, with quasi-degenerate ground states, low lying excited states, and open shell configurations.

A few full DC MCSCF implementations are already available.[19–21,48,49] In principle, a SFDC MCSCF calculation can be performed by using one of the existing DC implementations, for instance, with a quaternion-based DC implementation*, by zeroing the various

*Note that the results obtained by zeroing the quaternion-imaginary terms in a full DC calculation are not numerically identical to the ones obtained with a genuine SF approach, although the difference is very small. This difference lies in the fact that a quaternion-real contribution arising from the small component integrals contains the products of quaternion-imaginary quantities, which are absent in a purely SFDC approach

quaternion-imaginary terms.[50] Here, we propose for the first time an approach specifically intended for the SFDC Hamiltonian which fully exploits all the computational advantages offered by the spin separation. Our SFDC-CASSCF scheme allows one to incorporate scalar relativistic effects in a multireference description in a cost-effective, but rigorous way. Furthermore, in the perspective of using a CASSCF wave function as a starting point for a correlated treatment,[51–55] a SFDC approach confines the computational overhead due to relativity to the non-rate-determining CASSCF step.

Several implementations have been proposed so far for the CASSCF approach. The various strategies can be grouped in two main families. First-order implementations[56–60] tackle the optimization of the CI coefficients and of the MO separately and are based on the generalized Brillouin theorem.[61] From an optimization theory point of view, they are called first-order as the convergence rate is linear in the number of iterations. Such methods, however, do not guarantee convergence and the lack of general mathematical results force their implementations to be manually calibrated on test cases. The second family of implementations, second-order methods,[62–68] represented therefore a major advancement in MCSCF theory, as they allow for a much more robust convergence. Second-order optimization methods are quadratically convergent and take explicitly into account the coupling between the CI and MO degrees of freedom. We pursue here a second-order implementation and follow closely the norm-extended optimization (NEO) scheme by Jensen and coworkers.[64,65] Our choice is motivated by two main considerations. A second-order scheme is more expensive than a first order one but, on the other hand, allows for tighter and more robust convergence.[46] This is particularly beneficial in the perspective of coupling our SFDC CASSCF implementation to a correlated treatment such as, for instance, the one provided by internally contracted Coupled Cluster theory. Among the various second order schemes, the one proposed by Meyer, Werner and Knowles[63,68] (MWK) has been shown to be particularly fast and robust, especially for systems where the computational cost is dominated by the orbital optimization part. However, its extension to a four-component relativistic treatment,

which requires one to generalize it to a high-order saddle point optimization, does not appear to be straightforward. On the contrary, the NEO scheme is immediately extensible to the four-component relativistic case.[19] Its guaranteed convergence[64,69] and overall excellent performances[70] make it thus the most convenient choice for our scopes. A more detailed comparison between the NEO and the MWK algorithms can be found in ref. 70.

This paper is organized as follows. In section 2, the theoretical foundations of our implementation are presented. Section 3 describes the implementation of the SFDC-CASSCF method with a focus on the computational cost of the various steps involved in the calculation. In section 4, numerical results are presented and a pilot application is shown. Section 5 ends the paper with some conclusions and perspectives.

2 Theory

In this section, the general tools needed in order to achieve a quadratically convergent CASSCF implementation for the spin-free Dirac-Coulomb Hamiltonian are briefly presented. First, the spin separation of the DC Hamiltonian is discussed in section 2.1 and the spin-free Dirac Coulomb Hamiltonian is introduced. A parametrization suitable for the SFDC-CASSCF wavefunction, which is needed in order to define the optimization procedure, is then discussed in section 2.2, with a focus on the computational advantages offered by the use of a spin-free Hamiltonian. The second-order optimization scheme we adopt for our implementation is finally briefly summarized in section 2.3 and the issues specific to a four-component relativistic description are analyzed.

2.1 The spin-free Dirac-Coulomb Hamiltonian

In this section, the main ideas underlying the exact spin separation of the Dirac-Coulomb Hamiltonian as proposed by Dyall[41] are shortly summarized. Using the formalism of second quantization and atomic units, the DC Hamiltonian reads, in its standard (Dirac) represen-

tation,

$$\mathscr{H}^{\mathrm{DC}} = \sum_{PQ} h^{\mathrm{D}}_{PQ} \hat{a}^{\dagger}_P \hat{a}_Q + \frac{1}{2} \sum_{PQRS} G_{PQRS} \hat{a}^{\dagger}_P \hat{a}^{\dagger}_Q \hat{a}_S \hat{a}_R, \tag{1}$$

where we denote by $P, Q, \ldots$ one-electron Dirac four components spinors ψ_P, which can be expressed in terms of two two-component spinors usually referred to as large component ψ^L_P and small component ψ^S_P

$$|\psi_P\rangle = \begin{pmatrix} \psi^L_P \\ \psi^S_P \end{pmatrix}, \tag{2}$$

with

$$|\psi^L_P\rangle = \begin{pmatrix} \psi^{L,\alpha}_P \\ \psi^{L,\beta}_P \end{pmatrix}, \qquad |\psi^S_P\rangle = \begin{pmatrix} \psi^{S,\alpha}_P \\ \psi^{S,\beta}_P \end{pmatrix}. \tag{3}$$

h^{D} is the one-particle Dirac operator, whose matrix elements are

$$\begin{aligned} h^{\mathrm{D}}_{PQ} &= \left\langle \begin{pmatrix} \psi^L_P \\ \psi^S_P \end{pmatrix} \middle| \begin{pmatrix} V & c\vec{\sigma}\cdot\vec{p} \\ c\vec{\sigma}\cdot\vec{p} & V - 2c^2 \end{pmatrix} \middle| \begin{pmatrix} \psi^L_Q \\ \psi^S_Q \end{pmatrix} \right\rangle \\ &= \langle \psi^L_P | V | \psi^L_Q \rangle + c(\langle \psi^L_P | \vec{\sigma}\cdot\vec{p} | \psi^S_Q \rangle + \langle \psi^S_P | \vec{\sigma}\cdot\vec{p} | \psi^L_Q \rangle) + \langle \psi^S_P | V - 2c^2 | \psi^S_Q \rangle, \end{aligned} \tag{4}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ is a vector collecting the three Pauli matrices, $\vec{p}$ is the momentum operator, V is the nuclear electrostatic potential, c the speed of light and we have shifted the energy by $-c^2$ in order to match the non-relativistic energy scale. G is the (instantaneous) Coulomb interaction operator, whose matrix elements are

$$\begin{aligned} G_{PQRS} &= \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\psi_P(\mathbf{r}_1)^{\dagger} \psi_R(\mathbf{r}_1) \psi_Q(\mathbf{r}_2)^{\dagger} \psi_S(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} = (PR|QS) \\ &= (\psi^L_P \psi^L_R | \psi^L_Q \psi^L_S) + (\psi^L_P \psi^L_R | \psi^S_Q \psi^S_S) + (\psi^S_P \psi^S_R | \psi^L_Q \psi^L_S) + (\psi^S_P \psi^S_R | \psi^S_Q \psi^S_S), \end{aligned} \tag{5}$$

where we introduced Mulliken's notation for the electron repulsion integrals (ERIs).

In order to achieve spin separation we apply the following metric change to the small

component

$$\psi_P^S = \frac{\vec{\sigma}\cdot\vec{p}}{2c}\phi_P^L, \tag{6}$$

where ϕ_P^L is named the "pseudo large" component, as it possesses the same symmetries and, for positive energy solutions, the same order of magnitude as the large component. Furthermore, in the non-relativistic limit, it converges to the large component. The matrix elements of the Dirac-Coulomb Hamiltonian are then expressed in terms of the large and pseudo-large component, thus folding eq. 6 into the operator.[41] Finally, spin separation is achieved by using the Dirac identity

$$(\vec{\sigma}\cdot\vec{u})(\vec{\sigma}\cdot\vec{v}) = \vec{u}\cdot\vec{v} + i\vec{\sigma}\cdot(\vec{u}\times\vec{v}). \tag{7}$$

Both the one- and two-body operators decompose exactly in the sum of a spin-free and spin-orbit part

$$h_{PQ}^{\mathrm{D}} = h_{PQ}^{\mathrm{D,SF}} + h_{PQ}^{\mathrm{D,SO}}; \quad G_{PQRS} = G_{PQRS}^{\mathrm{SF}} + G_{PQRS}^{\mathrm{SO}}. \tag{8}$$

The spin-free Dirac-Coulomb Hamiltonian is obtained by neglecting the SO terms

$$\mathscr{H}^{\mathrm{SFDC}} = \sum_{PQ} h_{PQ}^{\mathrm{D,SF}} \hat{a}_P^\dagger \hat{a}_Q + \frac{1}{2}\sum_{PQRS} G_{PQRS}^{\mathrm{SF}} \hat{a}_P^\dagger \hat{a}_Q^\dagger \hat{a}_S \hat{a}_R, \tag{9}$$

where

$$h_{PQ}^{\mathrm{D,SF}} = \langle\psi_P^L|V|\psi_Q^L\rangle + \langle\psi_P^L|T|\phi_Q^L\rangle + \langle\phi_P^L|T|\psi_Q^L\rangle + \langle\phi_P^L|\frac{\vec{p}\cdot V\vec{p}}{4c^2} - T|\phi_Q^L\rangle, \tag{10}$$

$$\begin{aligned} G_{PQRS}^{\mathrm{SF}} &= (\psi_P^L\psi_R^L|\psi_Q^L\psi_S^L) + \frac{1}{4c^2}([\vec{p}\phi_P^L]\cdot[\vec{p}\phi_R^L]|\psi_Q^L\psi_S^L) \\ &\quad + \frac{1}{4c^2}(\psi_P^L\psi_R^L|[\vec{p}\phi_Q^L]\cdot[\vec{p}\phi_S^L]) + \frac{1}{16c^4}([\vec{p}\phi_P^L]\cdot[\vec{p}\phi_R^L]|[\vec{p}\phi_Q^L]\cdot[\vec{p}\phi_S^L]). \end{aligned} \tag{11}$$

The SF approximation introduces two major computational advantages that can be immediately deduced by looking at the SFDC Hamiltonian: neither $h_{PQ}^{\mathrm{D,SF}}$ nor G_{PQRS}^{SF} depends on

the spin operator, nor do they contain terms that require a complex representation. The SFDC Hamiltonian retains therefore full spatial and spin symmetry and a real wave function can be chosen. The same basis functions can be used to expand both the large and pseudo-large component (in particular, this is equivalent to using a restricted-kinetic-balanced basis set[71]). We can therefore write the SFDC Hamiltonian as

$$\mathscr{H}^{\mathrm{SFDC}} = \sum_{PQ}\sum_{\sigma} h_{PQ}^{\mathrm{D,SF}} \hat{a}_{P\sigma}^{\dagger}\hat{a}_{Q\sigma} + \frac{1}{2}\sum_{PQRS}\sum_{\sigma\tau} G_{PQRS}^{\mathrm{SF}} \hat{a}_{P\sigma}^{\dagger}\hat{a}_{R\tau}^{\dagger}\hat{a}_{S\tau}\hat{a}_{Q\sigma}, \tag{12}$$

where σ, τ are spin indices and the indices $P, Q, ...$ run over $2N_b$ orthonormal spatial orbitals, obtained for instance by solving the spin-free Dirac-Fock equations,[30] of which N_b are associated with positive energy states (PES) and N_b with negative energy states (NES). In order to simplify the notation, we introduce the operators

$$\hat{E}_{PQ} = \sum_{\sigma} \hat{a}_{P\sigma}^{\dagger}\hat{a}_{Q\sigma}, \quad \hat{e}_{PQRS} = \sum_{\sigma\tau} \hat{a}_{P\sigma}^{\dagger}\hat{a}_{R\tau}^{\dagger}\hat{a}_{S\tau}\hat{a}_{Q\sigma}. \tag{13}$$

Omitting the SFDC labels, the SFDC Hamiltonian reads

$$\mathscr{H} = \sum_{PQ} h_{PQ}\hat{E}_{PQ} + \frac{1}{2}\sum_{PQRS} (PQ|RS)\hat{e}_{PQRS}. \tag{14}$$

Thanks to our choice of real basis functions, all the matrix elements in eq. 14 are real. We stress again how crucial this is by pointing out that the Hamiltonian in eq. 14 is formally identical to a non-relativistic Hamiltonian which means that, by working in the MO basis and adopting the no-pair approximation,[72] any post-HF treatment can be performed exactly as in the non-relativistic case.[30]

2.2 Parametrization of the CASSCF wave function for the SFDC problem

In this section, we introduce the CASSCF wave function and a parametrization suitable for its variational optimization. We assume here that an initial guess for the MO coefficients is available, for instance, the solution to the SF Dirac-Hartree-Fock equations. Thanks to the properties of the SFDC Hamiltonian, the MOs can be chosen real. Furthermore, spin conservation allows us to adopt a restricted approach and to build our wavefunction starting from a set of $2N_b$ *spatial* molecular orbitals $\{\varphi_P\}_{P=1}^{2N_b}$, i.e., N_b positive energy states (PES) and N_b negative energy states (NES). In the following, we use the indices $p, q, \ldots$ to label a generic PES and $\tilde{p}, \tilde{q} \ldots$ to label NES. From now on, we use capital letters $P, Q, \ldots$ to label a generic one-particle function (and not a four-component spinor).

The CASSCF method divides the orbital space in three sets, which are used to determine which Slater determinants form the CASSCF wave function. Internal orbitals are doubly occupied in every determinant, external orbitals are empty in every determinant and active orbitals have no restriction on their occupation. For the relativistic case, a fourth set of orbitals has to be considered, i.e., the NES, which, as the external orbitals, are always unoccupied. The CASSCF ansatz is the linear combination of all the Slater determinants obtained by varying the occupation of the active orbitals. The set of MOs used to build the Slater determinants is optimized together with the coefficients of the determinants. A convenient, global parameterization of the CASSCF wave function, introduced by Jensen et al. as a starting point for their norm-extended optimization (NEO) algorithm,[65–67] is

$$|\Psi\rangle = e^{-\kappa} \frac{|0\rangle + \hat{P}|\mathbf{c}\rangle}{\||0\rangle + \hat{P}|\mathbf{c}\rangle\|}. \tag{15}$$

In eq. 15, $|0\rangle = \sum_{I=1}^{N\mathrm{det}} c_I^0 |\Phi_I\rangle$ is the current approximation to the CASSCF wave function, to which we refer as current expansion point (CEP), $\hat{P} = 1 - |0\rangle\langle 0|$ is the orthogonal projector

onto the orthogonal complement to $|0\rangle$ and

$$|\mathbf{c}\rangle = \sum_{I=1}^{N_{\text{det}}} c_I |\Phi_I\rangle$$

is a correction vector. A variation of the CI coefficients is therefore parametrized by adding a linear combination of determinants orthogonal to the CEP. As this does not preserve the wave function norm, normalization has to be carried out explicitly by dividing by the norm

$$\||0\rangle + \hat{P}|\mathbf{c}\rangle\| = \sqrt{1 + \langle \mathbf{c}|\hat{P}|\mathbf{c}\rangle}.$$

Spin symmetry plays an important role in the definition of the CASSCF CI space, as it allows one to exclude a large number of terms from the CI expansion. The total number of determinants is

$$N_{\text{det}} = \binom{N_{\text{act}}}{m_\alpha}\binom{N_{\text{act}}}{m_\beta}, \tag{16}$$

which is to be compared to $\binom{2N_{\text{act}}}{m_\alpha+m_\beta}$ for the full DC case, where m_α and m_β are the number of α and β electrons, respectively. This means that the size of the SFDC-CASSCF CI problem is the same as for its non-relativistic counterpart. As the full CI step dominates the computational cost of a CASSCF calculation for large enough active spaces, this is a major advantage of a SFDC approach. The effect of spin conservation is particularly dramatic for open-shell systems, as the number of allowed determinants is greatly reduced for high-spin molecules. As an example, we report in table 1 the number of determinants $N_{\text{det}}^{\text{SFDC}}$ needed in the SF case for accommodating 10 electrons in 10 active orbitals (a CAS(10,10) calculation), for various values of M_S. The ratio between the number of determinants required in a full DC calculation and $N_{\text{det}}^{\text{SFDC}}$, which does not depend on M_S, is also reported. The SFDC-CASSCF CI expansion coefficients can be chosen real and point group symmetry can be fully exploited by only considering the determinants that transform according the desired irreducible representation (irrep).

Table 1: Number of Slater determinants in the CI expansion for a CAS(10,10) calculation and ratio with respect to the number needed in a full DC calculation, which is always 184756.

M_S	N_{det}	ratio
0	63504	3
1	44100	4
2	14400	13
3	2025	91
4	100	1848
5	1	184756

Variations in the orbital coefficients are parametrized in eq. 15 by introducing a unitary transformation

$$\hat{U} = e^{-\hat{\kappa}}; \quad \hat{\kappa} = \sum_{P>Q} \kappa_{PQ} \left(\hat{E}_{PQ} - \hat{E}_{QP} \right) = \sum_{P>Q} \kappa_{PQ} \hat{E}^{-}_{PQ}. \tag{17}$$

Again, spin symmetry allows one to use a restricted approach, halving thus the size of the MO coefficients matrix and reducing the number of degrees of freedom needed to parametrize their variation. Point group symmetry can be used by enforcing that, as orbitals transforming according to different irreps do not mix, all κ_{PQ} coefficients with P and Q belonging to different irreps vanish. Finally, real orbitals can be rotated with an orthogonal transformation, i.e., the $\boldsymbol{\kappa}$ coefficients can be chosen real. Spin separation introduces therefore considerable simplifications in the orbital optimization problem with respect to the full DC case. However, orbital optimization is more complex for the SFDC case than for its non-relativistic counterpart because of the presence of the NES. The first, major source of overhead stems from the use of uncontracted basis sets, which is necessary when using restricted kinetic balanced basis sets in order to avoid variational collapse issues. This increases significantly the number of basis functions that one has to use with respect to what required for a non-relativistic computation, which in turn increases both the computational cost and the memory requirements of all the operations associated with orbital optimization. Furthermore, rotations between PES and NES have to be taken into account, increasing thus the number of variational

parameters. Finally, the SFDC-CASSCF energy needs to be *maximized* with respect to PES-NES rotations, so that the variational problem is not a minimization, but a high-order saddle point optimization. The orbital rotation operator mixes, in principle, every orbital with every other orbital. However, it is easily proven that some of these rotations are redundant, as they do not affect the energy. In particular, rotations mixing internal orbitals with internal orbitals, external with external or negative with negative do not affect the energy. Within the CASSCF ansatz, the active with active orbital rotations are also redundant. Finally, rotations between NES and external orbitals leave the energy unchanged, too. In conclusion, the (non-redundant) orbital rotation operator has the following expression

$$\hat{\kappa} = \sum_{i=1}^{N_{\text{int}}} \sum_{u=1}^{N_{\text{act}}} \kappa_{iu}^{+} \hat{E}_{iu}^{-} + \sum_{i=1}^{N_{\text{int}}} \sum_{a=1}^{N_{\text{ext}}} \kappa_{ia}^{+} \hat{E}_{ia}^{-} + \sum_{i=1}^{N_{\text{int}}} \sum_{\tilde{p}=1}^{N_{\text{neg}}} \kappa_{i\tilde{p}}^{-} \hat{E}_{i\tilde{p}}^{-} + \sum_{u=1}^{N_{\text{act}}} \sum_{a=1}^{N_{\text{ext}}} \kappa_{ua}^{+} \hat{E}_{ua}^{-} + \sum_{u=1}^{N_{\text{act}}} \sum_{\tilde{p}=1}^{N_{\text{neg}}} \kappa_{u\tilde{p}}^{-} \hat{E}_{u\tilde{p}}^{-}, \tag{18}$$

where we label with the superscript + the rotation parameters that correspond to PES-PES rotations and with a superscript − the ones that correspond to PES-NES rotations. We remark again that the latter rotations are characteristic of a four-component relativistic approach.

2.3 Optimization of the SFDC-CASSCF wave function

As already mentioned in section 2.2, a SF relativistic CASSCF calculation differs from a non-relativistic one not only because of the presence of the $\boldsymbol{\kappa}^{-}$ parameters, but also because the optimal parameters are determined by minimizing the energy with respect to $\mathbf{c}$ and $\boldsymbol{\kappa}^{+}$ and by maximizing it with respect to the $\boldsymbol{\kappa}^{-}$ rotations. A high-order saddle-point optimization problem has therefore to be solved. In this section, the optimization strategy, based on the NEO method,[65–67] will be presented and the issues arising from the relativistic saddle-point problem addressed. A complete derivation of the NEO method can be found in ref. 73.

In a second-order scheme, the energy is expanded in a Taylor series up to second or-

der around the CEP with respect to the parameters introduced in section 2.2. Let $\mathbf{x} = (\mathbf{c}, \boldsymbol{\kappa}^+, \boldsymbol{\kappa}^-)$ denote a point in the CASSCF parameter space, $\mathbf{x}_0 = (\mathbf{c}^0, 0, 0)$ be the CEP, and $\mathbf{P} = \mathbf{I} - \mathbf{x}_0\mathbf{x}_0^\dagger$ be the matrix representation of the projector $\hat{P}$. Expanding the CASSCF energy up to second order, one gets

$$\mathcal{E}(\mathbf{x}) \approx E_0 + \mathbf{g}^\dagger(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^\dagger \mathbf{G}(\mathbf{x} - \mathbf{x}_0) := \mathcal{Q}(\mathbf{x}), \tag{19}$$

where we have introduced the quadratic model $\mathcal{Q}(\mathbf{x})$, $E_0 = \mathcal{E}(\mathbf{x}_0)$, $\mathbf{g}$ is the gradient of the energy with respect to the variational parameters computed at the CEP

$$g_I^c = \frac{\partial \mathcal{E}}{\partial c_I} = 2\langle \Phi_I | \hat{P} \mathscr{H} | 0 \rangle \tag{20a}$$

$$g_{PQ}^o = \frac{\partial \mathcal{E}}{\partial \kappa_{PQ}} = \langle 0 | [\hat{E}_{PQ}^-, \mathscr{H}] | 0 \rangle \tag{20b}$$

and $\mathbf{G}$ is the Hessian computed at the CEP

$$\begin{aligned} G_{I,J}^{cc} &= \frac{\partial^2 \mathcal{E}}{\partial c_I \partial c_J} = 2\langle \Phi_I | \hat{P}(\mathscr{H} - E_0)\hat{P} | \Phi_J \rangle \\ G_{I,PQ}^{co} &= \frac{\partial^2 \mathcal{E}}{\partial c_I \partial \kappa_{PQ}} = 2\langle \Phi_I | \hat{P}[\hat{E}_{PQ}^-, \mathscr{H}] | 0 \rangle \\ G_{PQ,RS}^{oo} &= \frac{\partial^2 \mathcal{E}}{\partial \kappa_{PQ} \partial \kappa_{RS}} = \frac{1}{2}(1 + \hat{P}_{PQ,RS})\langle 0 | \hat{P}[\hat{E}_{PQ}^-, [\hat{E}_{RS}^-, \mathscr{H}]] | 0 \rangle. \end{aligned} \tag{21}$$

$\hat{P}_{PQ,RS}$ permutes the indices pairs PQ and RS. Note that the Hessian couples the variations of the CI coefficients to the ones of the orbitals and vice versa. Such a coupling is fundamental to have a truly second-order procedure. Working expressions for the energy, gradient and the product of the Hessian times a generic vector are provided in section 3. The simplest second-order method is the Newton-Raphson (NR) method, which looks for a stationary point of the quadratic model and uses it as a new expansion point in the following iteration. The NR method converges only if the starting point is close enough to the desired stationary

point. For this reason, the concept of *trust region* is introduced and the NR step

$$\boldsymbol{\delta} = -(\mathbf{PGP})^{-1}\mathbf{g}, \tag{22}$$

is accepted only if its norm is shorter than a user-defined trust radius R_t. Otherwise, a constrained search is performed, looking for a step to the border of the trust region. By introducing the constraint

$$\|\boldsymbol{\delta}\| = R_t \tag{23}$$

together with a Lagrange multiplier ν, the restricted step NR equations, also known as Levenberg-Marquardt (LM) equations,[69] are obtained:

$$(\mathbf{G} - \nu\mathbf{I})\boldsymbol{\delta} = -\mathbf{g} \tag{24a}$$

$$\|\boldsymbol{\delta}(\nu)\| = R_t. \tag{24b}$$

The NEO method is a computationally efficient and elegant method to solve the system of equations 24. This is done by introducing a gradient-scaled augmented Hessian

$$\mathbf{L}(\alpha) = \mathbf{G} + \alpha(\mathbf{x}_0\mathbf{g}^\dagger + \mathbf{g}\mathbf{x}_0^\dagger), \tag{25}$$

where α is a positive, real number. The NEO step is computed as

$$\boldsymbol{\delta} = \frac{1}{s_i\alpha}\mathbf{P}\mathbf{y}_i \tag{26}$$

where

$$\mathbf{L}(\alpha)\mathbf{y}_i = \lambda_i\mathbf{y}_i. \tag{27}$$

It can be shown[65] that the NEO step solves eq. 24 with a level shift parameter $\nu = \lambda_i$. The level shift parameter has to be chosen such that the level-shifted Hessian $\mathbf{G} - \lambda_i\mathbf{I}$ has

the desired signature. For a saddle-point optimization problem, which is of interest for the SFDC-CASSCF problem, we need

$$\mu_{N_{\mathrm{PN}}} < \lambda_i < \mu_{N_{\mathrm{PN}}+1}, \tag{28}$$

where $N_{\mathrm{PN}} = N_{\mathrm{neg}} N_{\mathrm{int}} + N_{\mathrm{neg}} N_{\mathrm{act}}$ is the number of PES-NES rotations. As the NEO matrix satisfies the Hylleraas-Undheim-MacDonald theorem,[74,75] i.e.,

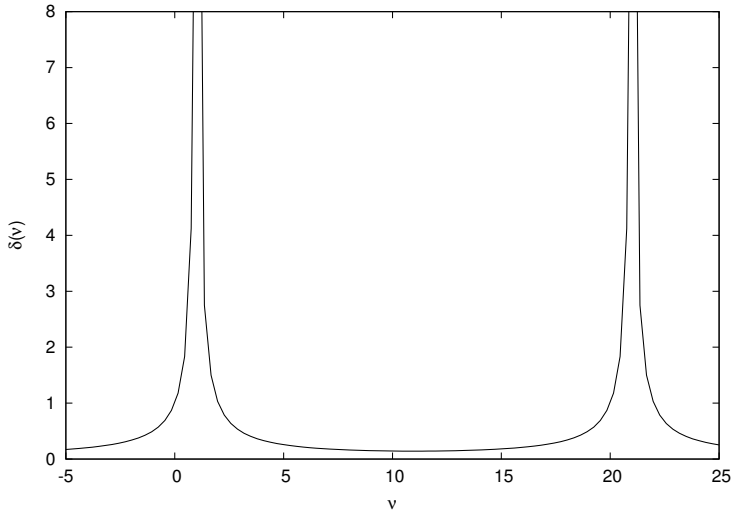
$$\mathbf{PL}(\alpha)\mathbf{P} = \mathbf{L}(0) = \mathbf{G},$$

the eigenvalues of the Hessian are interleaved between the eigenvalues of $\mathbf{L}(\alpha)$, which means that the $(N_{\mathrm{PN}} + 1)$-th eigenvalue of $\mathbf{L}$ automatically satisfies eq. 28. In order to enforce the restricted-step condition, a solution to eq. 24b needs to be found. While a solution to such an equation can always be found for $\nu < \mu_1$ and $\nu > \mu_n$, i.e., for minimization and maximization problems, this is not in general the case for saddle points. It is possible to show that, in the interval defined by eq. 28, $\|\delta(\nu)\|$ is a positive function which goes to infinity for ν approaching an eigenvalues of the Hessian and which has a unique, positive minimum. A solution to eq. 24b can therefore be found only if such a minimum is smaller than R_t. A rough estimate of the value of the minimum between two eigenvalues μ_i and μ_{i+1} is given by a constant times $(\mu_{i+1} - \mu_i)^{-1}$.[69] Therefore, if the two eigenvalues that define the saddle point we are interested in are well separated, finding a solution to eq. 24b is not a problem in practice. This is fortunately the case, as $\mu_{N_{\mathrm{PN}+1}} - \mu_{N_{\mathrm{PN}}} \approx 2mc^2$. A pictorial representation of the behavior of $\|\delta(\nu)\|$ in such a case is shown in figure 1. In practice, the norm of the NEO step depends on the parameter α introduced in eq. 25. To solve eq. 24a, a root of the function

$$f(\alpha) = \|\boldsymbol{\delta}(\alpha)\|^2 - R_t^2 \tag{29}$$

is found numerically. The choice of the trust radius is an important aspect of the optimization

Figure 1: Norm of the Levenberg-Marquardt step $\boldsymbol{\delta}(\nu)$ as a function of the level shift parameter ν for a simple example. The function goes from zero to infinity for $\nu < \lambda_1$ and $\nu > \lambda_2$, which means that a solution to eq. 24a can always be found. The function has a positive minimum for $\lambda_1 < \nu < \lambda_2$, which is the interval to be targeted for a saddle-point optimization. However, such a minimum is close to zero if the eigenvalues are well separated, so that in practice, a solution to eq. 24a can be found if the eigenvalues are distant enough. This is the case for relativistic calculations, where the interesting saddle point lies between two eigenvalues of the Hessian which are roughly $2c^2 \approx 37500$ a. u. apart.



procedure. The concept of trust region originates from the limited agreement between the quadratic model of the energy and the energy itself. This idea was formalized by Fletcher,[69] who used it to define an adaptive trust radius by evaluating, at each step of the optimization procedure, the ratio between the predict change in the energy and the actual one, i.e., by defining

$$r = \frac{\mathcal{E}(\mathbf{x}) - \mathcal{E}_0}{\mathcal{Q}(\mathbf{x}) - E_0}. \tag{30}$$

The closer r is to one, the better the agreement. Fletcher suggests to update the trust radius as follows:

- if $r < 0$ reject the step, reduce the trust radius from R_t to aR_t, $a < 1$ and restart
- if $0 < r < 0.25$ accept the step, but still reduce the trust radius from R_t to aR_t
- if $0.25 < r < 0.75$ accept the step and leave the trust radius unchanged

- if $r > 0.75$ accept the step and increase the trust radius from R_t to bR_t, $b > 1$.

Fletcher claims that the method is not sensitive to the specific values of a and b, which can be chosen depending on the specific problem. In our implementation, we use $a = 0.66$ and $b = 1.2$. In any case, R_t is not allowed to be larger than 1.2 or smaller than 0.1. An important result holds: the Fletcher-LM method *always* converges to the closest minimum for a regular enough function.

We conclude this section by summarizing the main characteristics of the SFDC-CASSCF problem. With respect to a non-relativistic CASSCF, there are two main differences that are due to the presence of the NES. First, more orbital rotations need to be taken into account, which means that the number of parameters needed in the orbitals is increased (roughly, doubled). Second, such rotations complicate the wave function optimization which needs to aim at a high-order saddle point and not at a minimum. The CI problem which, we remark, is the cost-determining part of a CASSCF calculation is identical to the non-relativistic CASSCF-CI problem. On the contrary, a full DC-CASSCF implementation implies a much more computationally demanding CI problem, as the lack of spin symmetry increases the number of determinants that need to be included in the wave function. Complex algebra and the use of double groups instead of point-group symmetry further increase the cost of a full DC-CASSCF treatment.

3 Implementation

In this section, the implementation of SFDC-CASSCF method described in section 2 in the CFOUR[76] suite of programs is presented. The infrastructure for SFDC calculations present in CFOUR,[30] and in particular the relativistic integrals, constitutes the starting point of our implementation. First, working equations for the energy and gradient are given and the additional terms stemming from a relativistic treatment are underlined. An efficient and robust procedure to solve the NEO eigenvalue problem is then presented, together with

a strategy to compute the $(N_{\mathrm{PN}}+1)$-th eigenvalue which is needed for the saddle-point optimization.

3.1 Energy, gradient, and linear transformations

The NEO procedure requires at each iteration the computation of the current energy and gradient. The energy can be obtained as the expectation value of the Hamiltonian

$$E_0 = \langle 0|\mathscr{H}|0\rangle = \sum_{PQ} h_{PQ}\gamma_{PQ} + \frac{1}{2}\sum_{PQRS}(PQ|RS)\Gamma_{PQRS}, \tag{31}$$

where we have introduced the one- and two-body reduced density matrices (1- and 2-RDM)

$$\gamma_{PQ} = \langle 0|\hat{E}_{PQ}|0\rangle, \quad \Gamma_{PQRS} = \langle 0|\hat{e}_{PQRS}|0\rangle. \tag{32}$$

There are five classes of MO rotations with respect to which a MO gradient contribution has to be assembled, as explicitly pointed out in eq. 17. Three are standard PES-PES rotations, which are common to non-relativistic implementations

$$g_{iu} = \frac{\partial\mathcal{E}}{\partial\kappa^{+}_{iu}} = 2(F_{iu} - F_{ui}), \quad g_{ia} = \frac{\partial\mathcal{E}}{\partial\kappa^{+}_{ia}} = 2F_{ia}, \qquad g_{ua} = \frac{\partial\mathcal{E}}{\partial\kappa^{+}_{ua}} = 2F_{ua}. \tag{33}$$

The other two terms

$$g_{i\tilde{p}} = \frac{\partial\mathcal{E}}{\partial\kappa^{-}_{i\tilde{p}}} = 2F_{i\tilde{p}}, \quad g_{u\tilde{p}} = \frac{\partial\mathcal{E}}{\partial\kappa^{-}_{u\tilde{p}}} = 2F_{u\tilde{p}} \tag{34}$$

correspond to PES-NES rotations and are peculiar to four-component relativistic calculations. In eqs. 33 and 34 we have introduced the generalized Fock matrix F_{PQ}, whose elements are given by

$$F_{iP} = 2(F^{I}_{iP} + F^{A}_{iP}), \quad F_{uP} = \sum_{u}\gamma_{uv}F^{I}_{vP} + \sum_{vxy}\Gamma_{uvxy}(Pv|xy), \quad F_{aP} = F_{\tilde{p}P} = 0, \tag{35}$$

where

$$F^I_{PQ} = h_{PQ} + \sum_i \left(2(PQ|ii) - (Pi|Qi)\right), \tag{36}$$

$$F^A_{PQ} = \sum_{uv} \gamma_{uv} \left((PQ|uv) - \frac{1}{2}(Pv|Qu) \right). \tag{37}$$

Note that only the active 1- and 2-RDM, i.e., the 1- and 2-RDM with all indices running over the active orbitals manifold, are referenced. A four-component relativistic treatment has two main effects on the evaluation of the MO rotation gradient, which both concern the transformation of the electron repulsion integrals from the atomic orbitals (AO) to the MO basis. First, integrals involving the pseudo-large component need to be taken into account. This requires one to accumulate the transformed integrals as the result of four subsequent integral transformations, involving (LL|LL), (LL|SS), (SS|LL), (SS|SS) integrals, respectively, where L stands for large component and S for pseudo-large, with the pseudo-large component representing here the small component. Second, transformed integrals with indices belonging to the NES are required to assemble the PES-NES gradient contributions in eq. 34. With respect to the non-relativistic counterpart, the integral transformations involved in a SFDC-CASSCF calculation are roughly 10 times more expensive. Nevertheless, thanks to spin separation, the integral transformation is significantly less expensive than in a full DC calculation, as the number of independent integrals can be reduced by using spin symmetry and the possibility of using a restricted approach. Furthermore, full permutation symmetry is retained, real algebra is used, and point group symmetry can be fully exploited.

The CI gradient

$$g_I = \frac{\partial \mathcal{E}}{\partial c_I} = 2\langle \Phi_I | \sum_{uv} F^I_{uv} \hat{E}_{uv} + \frac{1}{2} \sum_{uvxy} \hat{e}_{uvxy} |0\rangle - 2c^0_I (E_0 - E_I), \tag{38}$$

where

$$E_I = \sum_i \left(h_{ii} + F^I_{ii} \right), \tag{39}$$

is assembled by computing a direct-CI step, i.e., by computing the action of the Hamiltonian

on $|0\rangle$. Note that only indices of active orbitals are referenced in eq. 38. The computation requires the active portion of F^I and MO integrals with all indices referring to active orbitals, which are already available from the MO gradient computation. As we already remarked in section 2, there is no difference in the CI step between a SFDC and a non-relativistic CASSCF calculation. The same routines can be used in both cases, the same number of CI coefficients is required and the same symmetries can be exploited to increase the efficiency of the computation. As the CI part is the rate-determining step of a CASSCF computation, the SFDC approximation introduces an important computational gain with respect to a full DC calculation. Our direct CI implementation uses Handy's α and β strings formalism,[77] therefore adopting a determinant CI strategy,[78,79] and follows the vector implementation suggested by Bendazzoli and Evangelisti[80] using a direct-list algorithm as described by Gagliardi et al.[81]

We conclude this section by reporting the expressions needed to compute the product of the NEO matrix with a trial vector, which are needed in order to solve iteratively the NEO equations. The NEO matrix, given in eq. 25, consists of the Hessian plus a rank-1 contribution. It has therefore the same block structure as the Hessian

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}^{cc} & \mathbf{L}^{co} \\ \mathbf{L}^{oc} & \mathbf{L}^{oo} \end{pmatrix}, \tag{40}$$

where the superscript are consistent with the ones in eq. 21. In our implementation, we assemble explicitly the *oo* block. Working expressions for such a block were given for the first time by Siegbahn et al. in ref. 62 and are reported in supporting information. For the other blocks, the direct approach proposed by Jensen et al.[67] is pursued. We report here the

final expressions for the *cc*, *co* and *oc* blocks:

$$\sum_J L_{I,J} v_J = 2(\langle \Phi_I | \mathscr{H} | \mathbf{v}_c \rangle - E_0 v_I) + (\alpha - 1) \left[c_I^0 \left(\sum_J v_J g_J \right) + g_I \left(\sum_J v_J x_J \right) \right] \quad (41)$$

$$\sum_{PQ} L_{I,PQ} v_{PQ} = 2\langle 0 | \tilde{\mathscr{H}} | \Phi_I \rangle + (\alpha - 2) c_I^0 \left(\sum_{PQ} g_{PQ} v_{RS} \right) \quad (42)$$

$$\sum_I L_{PQ,I} v_I = 2 g_{PQ}^T + (\alpha - 2) g_{PQ} \left(\sum_I c_I^0 v_J \right), \quad (43)$$

where $|\mathbf{v}_c\rangle = \sum_I v_{c,I} |\Phi_I\rangle$. In eq. 42 we have introduced a one-index transformed Hamiltonian

$$\tilde{\mathscr{H}} = \sum_{uv} \tilde{h}_{uv} \hat{E}_{uv} + \frac{1}{2} \sum_{uvxy} \tilde{g}_{uvxy} \hat{e}_{uvxy}, \quad (44)$$

where

$$\tilde{h}_{uv} = \sum_P (v_{uP} h_{Pv} + v_{vP} h_{uQ}) \quad (45)$$

and

$$\tilde{g}_{uvxy} = \sum_P (v_{uP}(Pv|xy) + v_{vP}(uP|xy) + v_{xP}(uv|Py) + v_{yP}(uv|xP)) . \quad (46)$$

In eq. 43, we have introduced a transition orbital gradient $\mathbf{g}^T$, which is computed exactly as described at the beginning of this section, but using the transition density matrices

$$\gamma_{uv}^T = \langle 0 | \hat{E}_{uv} | \mathbf{v} \rangle + \langle \mathbf{v} | \hat{E}_{uv} | 0 \rangle \quad (47)$$

$$\Gamma_{uvxy}^T = \langle 0 | \hat{e}_{uvxy} | \mathbf{v} \rangle + \langle \mathbf{v} | \hat{e}_{uvxy} | 0 \rangle. \quad (48)$$

to assemble the generalized Fock matrix.

3.2 Macro and microiterations

In this section, the iterative optimization procedure and the iterative diagonalization method used to solve the NEO equations are described. The CASSCF program requires as an input,

besides the definition of the active space, a set of starting MO coefficients and a guess for the CI expansion coefficients. The main iterative procedure proceeds as follows:

1. Set the main iteration (macroiteration) counter $k = 1$. Read or assemble an inital guess (the HF wave function is currently used in our implementation).

2. Compute the 1- and 2-body reduced density matrices (1- and 2-RDM) γ_{uv} and Γ_{uvxy} with all indices active.

3. Perform a partial integral transformation to assemble the $(PQ|Ri)$ and $(PQ|Ru)$ integrals.

4. Compute the generalized Fock matrix. Compute the energy and assemble the MO rotation gradient and Hessian. These operations are performed just after the partial integral transformations in order to avoid unnecessary disk I/O.

5. If $k > 1$, compute the ratio as in eq. 30 and check whether the step computed in the previous iteration is acceptable. If not, reduce the trust radius, compute a new step and go back to step 2, otherwise, adjust the trust radius according to Fletcher's prescriptions. The trust radius equation is solved numerically using Brent's method,[82] all the details on how to compute the step in practice can be found in ref. 67.

6. Compute the CI gradient (eq. 38).

7. Check for convergence. If the CAS-SCF wavefunction has converged, exit.

8. Iteratively solve the NEO equations (microiterations) and compute the step.

9. Rotate the MO coefficients and update the CEP. Increment the iteration count $k := k + 1$.

Convergence is achieved when the root-mean-square norm of both the MO and CI gradient is smaller than a user-given threshold ϵ.

The heart of the quadratically convergent CASSCF implementation is step 8, i.e., the iterative solution to the NEO equations in order to compute the step that solves the LM equations. The CI and mixed blocks of the NEO matrix can be very large, but they are sparse and the CI block is diagonally dominant. On the other hand, the MO block is dense, usually not diagonally dominant and it can be very ill-conditioned, with its eigenvalues spanning many orders of magnitude. Furthermore, we are interested in the $(N_{\mathrm{PN}}+1)$-th eigenvalue and not in the lowest. Our procedure to solve the NEO equations is based on the Davidson diagonalization[83,84] and involves a further correction proposed by Olsen.[85] In order to select the proper eigenvalue, we divide the test vectors in three types: $\mathbf{v}_c, \mathbf{v}_\kappa$ and $\mathbf{v}_{\tilde{\kappa}}$, i.e., purely CI vectors, purely PES/PES and purely PES/NES orbital rotation vectors, respectively. When a $\mathbf{v}_{\tilde{\kappa}}$ vector is added, a large, negative eigenvalue appears in the reduced matrix. To converge towards the $(N_{PN}+1)$-th eigenvalue it is therefore sufficient to look for the $(N_{\tilde{\kappa}}+1)$-th eigenvalue and eigenvector of the reduced matrix, where $N_{\tilde{\kappa}}$ is the number of $\mathbf{v}_{\tilde{\kappa}}$ vectors in the expansion subspace.[19] This choice for the test vectors allows one to concentrate the effort on the difficult part of the problem, i.e., the one associated with orbital rotations.

Davidson's method is particularly efficient for diagonally dominant matrices. Unfortunately, because of the MO rotation Hessian block of the NEO matrix, this is not the case for CASSCF and convergence of the diagonalization procedure can be slow. Jensen et al. suggested[70] to use localized orbitals in order to increase the diagonal dominant character of the Hessian. Here, we propose two different strategies. In the Davidson method, a new expansion vector $\tilde{\mathbf{v}}$ is computed as follows

$$\tilde{\mathbf{v}} = (\mathbf{D} - \lambda_k \mathbf{I})^{-1}\mathbf{r}, \tag{49}$$

where $\mathbf{D}$ is the diagonal of $\mathbf{L}$, λ_k is the current approximation to the desired eigenvalue and $\mathbf{r}$ the residual of the current approximation to the eigenvector $\mathbf{u}$. Our first strategy is a preconditioned Davidson diagonalization, where we replace $\mathbf{D}$ with a sparse approximation

to $\mathbf{L}$:

$$M_{ij} = \begin{cases} L_{ij} & |L_{ij}| > \text{tol or } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{50}$$

The linear system

$$(\mathbf{M} - \lambda_k \mathbf{I})\tilde{\mathbf{v}} = \mathbf{r} \tag{51}$$

is then solved by using the sparse LU decomposition[86] implementation by Saunders et al.[87] In eq. 50, we use a threshold tol of 0.5, which we lower to 0.1 close to convergence. Standard Davidson updates are used for CI test vectors. Such a strategy is efficient, as $\mathbf{M}$ is indeed sparse and its inversion can be achieved with a very limited computational cost, and effective, allowing for robust convergence in a limited number of iterations. However, the use of the "ideal" preconditioner, i.e., the $(\mathbf{L} - \lambda_k \mathbf{I})^{-1}$ matrix, would have the paradoxical result of making the iterative procedure stagnate.[88] Therefore, the preconditioner should be good, but not too good. This evanescent statement makes a rigorous approach to preconditioning difficult and requires a case-specific calibration. Our choices for the $\mathbf{M}$ matrix reflect this kind of approach.

A completely different strategy, which avoids such ambiguity in the preconditioned Davidson method, is offered by the Jacobi-Davidson[88] (JD) method. JD replaces the Davidson update by a vector which is orthogonal to the current approximation $\mathbf{u}$ and solves the following linear system:

$$(\mathbf{I} - \mathbf{u}\mathbf{u}^\dagger)(\mathbf{L} - \lambda_k \mathbf{I})(1 - \mathbf{u}\mathbf{u}^\dagger)\tilde{\mathbf{v}} = -\mathbf{r}. \tag{52}$$

If the linear system in eq. 52 is solved exactly, the JD method is equivalent to the Rayleigh-Quotient inverse iteration,[88] which exhibits cubic convergence. Superlinear convergence can still be achieved by solving eq. 52 approximately with a (preconditioned) Krylov subspace solver for the linear system, such as GMRES.[89] Albeit more expensive than the incomplete LU preconditioned Davidson scheme, the Jacobi-Davidson method has the advantage of moving the computational effort to the solution of a linear system of equations, for which

preconditioning techniques are well established and iterative methods can be considered almost black box. In our implementation, we use the LU decomposition of a sparse approximation to $\mathbf{L}$ as a preconditioner, as in eq. 50. However, we use a tighter threshold of 0.01 to select the matrix elements of $\mathbf{M}$. While this makes the sparse LU decomposition more expensive, the preconditioner does not change during the iterative solution of eq. 52, so the LU decomposition only has to be computed once per JD iteration. As the convergence of the JD scheme depends on the quality of the guess for the eigenvalue, it is a good strategy to perform a few initial Davidson iterations and switch to JD only when a preliminary convergence criterion has been met.

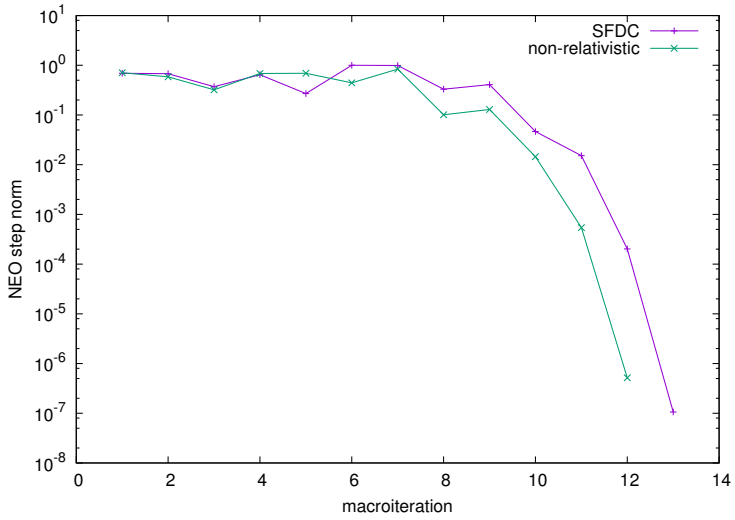
4 Numerical results

In this section, we present some numerical results obtained with our SFDC-CASSCF implementation. We start by showing the convergence behavior of the second-order optimization scheme and of the microiteration solvers described in section 3. As a pilot application, we computed the equilibrium geometry of the unsaturated monocarbonyls of nickel, palladium and platinum using both our non-relativistic and SFDC-CASSCF implementation.

4.1 Convergence properties of the macro and microiterations

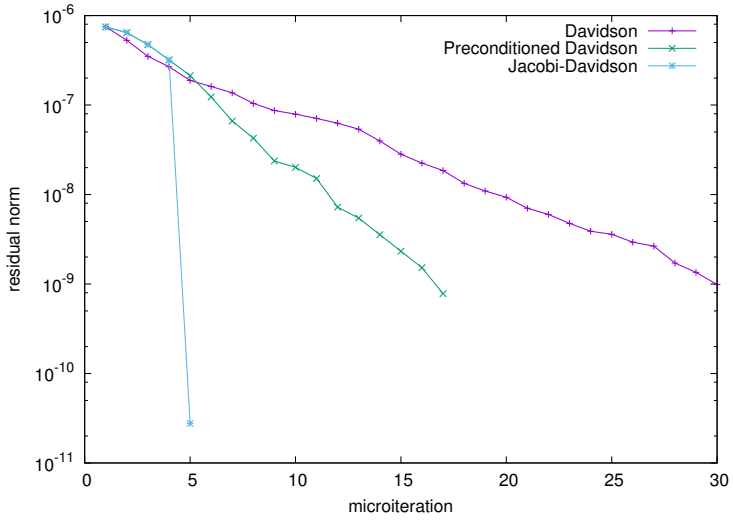
The NEO scheme's quadratic convergence properties have been formally proved by Jensen et al.[65] As an example of the convergence behavior of the NEO scheme for a SFDC calculation, we report in figure 2 the convergence profile of both a SFDC and a non-relativistic CASSCF calculation for the ${}^7\Pi_u$ state of $Cr(CO)_2$. The uncontracted cc-pVTZ[90,91] basis set was used. 10 electrons (8 α and 2 β) were correlated in an active space consisting of 12 orbitals as suggested by Kim et al.[92] The Cr-C and C-O distances were taken from the same reference and are, respectively, 2.077 Å and 1.110 Å. A tight convergence threshold of 10^{-10} in the RMS gradient was required. The SFDC calculation exhibits a convergence

Figure 2: Convergence profile of both a SFDC and a non-relativistic CASSCF calculation for the $^7\Pi_u$ state of $Cr(CO)_2$ correlating 10 electrons in 12 orbitals and using the uncontracted cc-pVTZ basis set. The NEO step norm is reported as a function of the number of macroiterations.



behavior very similar to its non-relativistic counterpart. In both cases, the first optimization steps are required to search for the quadratic region, the constraint on the step norm is active and large energy changes are observed. The norm of the step oscillates, but remains large and constant. Once the quadratic region is reached, convergence is rapidly achieved in a few iterations. The relativistic calculation requires one macroiteration more than the non-relativistic one, which is expected due the larger number of degrees of freedom involved in the SFDC optimization. In our experience, the example here reported reproduces the typical behavior of a SFDC-CASSCF calculation, both in terms of convergence pattern and of number of iterations required to achieve convergence as compared to the non-relativistic case. In order to show the convergence behavior of the microiteration solvers, we focus on the 12th iteration of the example reported in figure 2 for the SFDC calculation and report in figure 3 the convergence pattern observed by using a standard Davidson diagonalization, the preconditioned Davidson introduced in section 4, and the JD method. The convergence threshold for the microiterations is chosen depending on the norm of the gradient so that the quadratic convergence of the macroiterations is retained. In particular, it is fixed at $0.1\|\mathbf{g}\|^2$.

Figure 3: Convergence profile for the macroiteration associated with the last macroiteration for the SFDC-CASSCF calculation on $Cr(CO)_2$ using a (10,12) active space and the uncontracted cc-pVTZ basis set. The norm of the residual is reported as a function of the number of iteration.



For the case shown here, the convergence criterion was set to 10^{-9}. It is apparent from fig. 3 that the preconditioned Davidson scheme converges much quicker than the original and unmodified Davidson scheme, as it needs roughly one half of the iterations to compute an eigenvector with the required precision. The JD calculation shares the first four iterations with the preconditioned Davidson scheme. This is due to an implementation choice, as we compute four steps using the preconditioned Davidson scheme before switching to JD in order to provide a more stable guess. As a consequence, figure 3 contains one single JD iteration, which, due to the superlinear convergence of JD, is sufficient to reduce the residual to far less than the required threshold. The impressive efficiency of JD comes at a price, as in order to solve the linear system in eq. 52, 15 iterations of GMRES were needed. As a consequence, the JD scheme is slightly more expensive than the preconditioned Davidson as it requires more matrix-vector multiplications. Nevertheless, the more rigorous grounds on which the JD scheme is based render it a viable and not-too-expensive alternative for difficult cases. As a final remark, we point out that the convergence of the microiterations for the SFDC case is similar to the one observed for the non-relativistic case. The additional

PES-NES rotation are well decoupled from the PES-PES rotations and the CI coefficients, and usually only one or two $\mathbf{v}_{\tilde{\kappa}}$ trial vectors are sufficient to achieve convergence. As a consequence, no significant difference in the number of iterations is observed between a SFDC and a non-relativistic CASSCF calculations.

4.2 Structural properties of the monocarbonyls of the nickel group

Transition metal unsaturated monocarbonyls represent a good test case for our SFDC-CASSCF implementation, as they exhibit both multireference character and, for heavy transition metals, pronounced scalar relativistic effects. In particular, we aim at reproducing the trends in the metal-carbon distance and stretching frequency within the nickel group. Nickel,[93–98] palladium[99–101] and platinum[102–104] monocarbonyls have been extensively studied in the literature and both theoretical and experimental results are available.[105] Quantitative results have been obtained on the structures and electronic properties of such compounds,[93–97,99,102] especially on NiCO.[106] Here, we focus mainly on the impact of scalar relativistic effects on the structure and properties of the metal monocarbonyls, which are determined by comparing SFDC and non-relativistic results.

The uncontracted ANO-RCC basis set[107] was used for all systems. The CO molecule acts as a σ-donor, π-acceptor ligand, so we included in the active space the valence σ and π orbitals, together with a weakly occupied σ orbital and two π antibonding orbitals. The δ orbitals which originate from the metal's $d_{x^2-y^2}$ and d_{xy} orbitals have a non-bonding character and were not included in the active space, as suggested by Xu et al.[97] In total, 8 electrons were correlated in 7 orbitals. Working in C_{2v} symmetry, this corresponds to 3 a_1, 2 b_1 and 2 b_2 orbitals. Numerical differentiation was used to compute the gradients for geometry optimizations, double numerical differentiation for the harmonic frequencies. The ground state of these carbonyls is a linear $^1\Sigma^+$ state. Among the three carbonyls, NiCO has the most pronounced multireference character. The principal configurations are the HF one (weight 0.89), the four obtained by exciting one electron from a π to a π^*

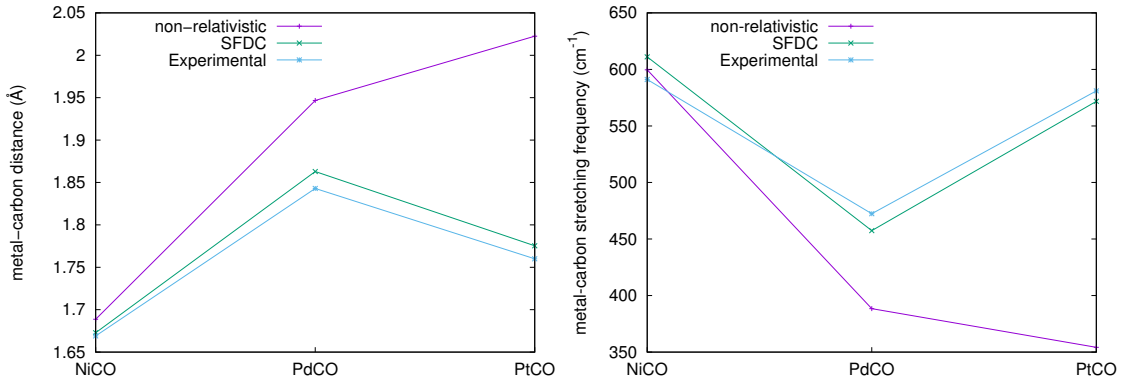
orbital (weight 0.15) and the two obtained exciting a σ electron to the low-lying σ virtual orbital (weight 0.15). We report in table 2 the geometrical parameters and the metal-carbon stretching harmonic frequencies obtained by both the SFDC and the non-relativistic CASSCF calculations. Experimental results are reported for comparison. To better compare

Table 2: Geometrical parameters (r_{MC}: metal-carbon distance, r_{CO}: carbon-oxygen distance, both in Å) and harmonic metal-carbon stretching frequencies (ω_{MC}, in cm^{-1}) for the monocarbonyl. The gradients for geometry optimizations and the harmonic frequencies were computed via numerical derivatives. A CAS(7,8) wave function and the uncontracted ANO-RCC basis set were used for all the calculations. The SFDC and non-relativistic (NR) CASSCF results are reported together with the experimental (Exp) data. Experimental data are taken from ref. 98 for NiCO, from refs. 100,101 for PdCO and from refs. 103,104 for PtCO.

	NiCO			PdCO			PtCO		
	NR	SFDC	Exp	NR	SFDC	Exp	NR	SFDC	Exp
r_{MC}	1.6887	1.6728	1.669	1.9467	1.8629	1.843	2.0225	1.7754	1.760
r_{CO}	1.1292	1.1300	1.153	1.1144	1.1180	1.140	1.1129	1.1255	1.146
ω_{MC}	600	611	591	388	457	472	354	572	581

the trends, the metal-carbon distances (left panel) and stretching frequencies (right panel) are also reported in figure 4. The CASSCF results are not expected to be in a perfect

Figure 4: Metal-carbon distances and harmonic stretching frequencies for NiCO, PdCO, and PtCO. Non-relativistic and SFDC CASSCF results were obtained using a CAS(7,8) wavefunction and the uncontracted ANO-RCC basis set.



agreement with the experimental ones, due to the lack of dynamic correlation treatment. Nevertheless, a semi-quantitative agreement can be observed if the CASSCF wave function

is flexible enough to correctly represent the molecular electronic ground state. The results for nickel, where relativistic effects are expected to be small, confirm that the CASSCF scheme provides a satisfactory description. Both the SFDC and the non-relativistic results are in good agreement with the experimental data. The comparison of SFDC and non-relativistic results is more interesting for the heavier PdCO and PtCO. In particular, comparing the same level of theory allows one to attribute the discrepancies in the results entirely to the relativistic treatment and therefore to better understand the quantitative effects of relativity on the structure and properties of a system. It is particularly interesting to compare the trends in both the metal-carbon distance and in the metal-carbon stretching frequency. Moving from nickel to palladium, an increase in the distance is observed, due to the larger covalent radius of the metal atom. However, such an increase is partly compensated by the contraction of the electronic shells in palladium, which is due to scalar relativistic effects. The discrepancy between the non-relativistic and SFDC results is therefore due to the fact that while the non-relativistic description only reproduces the first effect, a relativistic treatment is required to correctly describe the latter. The same observations are made for platinum, were the relativistic contraction effect is particularly strong. The non-relativistic description gives a metal-carbon bond length which is larger than the one for palladium and completely fails to reproduce the experimental trend, which shows a Pt-C distance shorter than the Pd-C one. On the contrary, the SFDC-CASSCF description correctly captures the contraction effect and gives a Pt-C bond length in good agreement with the experimental value. A very similar trend can be observed for the metal-carbon stretching frequency, which exhibits, as expected the opposite trend than the metal-carbon distance. In particular, it is lower for palladium than for both nickel and platinum. Again, a non-relativistic treatment fails to capture such a trend and gives a qualitatively wrong description. On the contrary, the SFDC-CASSCF treatment successfully reproduces the correct behavior, giving results in qualitative agreement with the experimental data.

5 Conclusions

In this article we have presented an implementation of the CASSCF method specifically tailored for a scalar relativistic treatment based on the spin-free Dirac-Coulomb Hamiltonian. Our implementation fully exploits all the computational advantages offered by the spin-separation of the DC Hamiltonian, namely, real algebra, spin symmetry, and point-group symmetry. This results in a method with a computational cost close to the one of a non-relativistic treatment, especially when large active spaces are used. The quadratic convergence of the scheme and an innovative numerical strategy to solve the optimization equations have been documented through an example. The applicability of the SFDC-CASSCF method has been demonstrated in a first application on the unsaturated monocarbonyls of the nickel group metals. It has been shown that the inclusion of SF relativistic effects is sufficient to correctly describe the behavior of the geometric parameters through the group and how such a behavior is mainly due to relativistic effects.

There are two main ingredients that are still lacking in our approach in order to achieve quantitative results. The first one is the inclusion of dynamic electronic correlation, using the CASSCF wavefunction as a starting point for a correlated treatment. A SFDC approach is particularly convenient for such a task, as no overhead is introduced in the rate-determining correlated post-treatment due to relativity. The second ingredient is the treatment of SO interactions. A perturbative treatment seams appropriate for not-too-heavy elements, such as the ones belonging to the first 5 rows of the periodic tables.[30] This direction is currently being explored in our group.

Supporting Information

Explicit expressions for the MO rotation Hessian. The Supporting Information is available free of charge on the ACS Publications website http://pubs.acs.org

Acknowledgments

F. L. would like to thank Professor Hans Jørgen Aagard Jensen for useful discussion on the norm-extended optimization scheme. The authors would also like to thank Professor Lan Cheng for carefully reading the manuscript. F. L. is grateful to the Alexander von Humboldt foundation for financial support.

References

(1) Pyykkö, P. *Chem. Rev.* **1988**, *88*, 563–594.

(2) Visscher, L. *J. Comput. Chem.* **2002**, *23*, 759–766.

(3) Dyall, K. G.; Fægri Jr., K. *Introduction to Relativistic Quantum Chemistry*; Oxford University Press, New York, 2007; Chapter Part IV.

(4) Liu, W. *Mol. Phys.* **2010**, *108*, 1679–1706.

(5) Saue, T. *ChemPhysChem* **2011**, *12*, 3077–3094.

(6) Fleig, T. *Chem. Phys.* **2012**, *395*, 2–15.

(7) Kutzelnigg, W. *Chem. Phys.* **2012**, *395*, 16–34.

(8) Liu, W. *Phys. Chem. Chem. Phys.* **2012**, *14*, 35–48.

(9) Cheng, L.; Stopkowicz, S.; Gauss, J. *Int. J. Quantum Chem.* **2014**, *114*, 1108–1127.

(10) Breit, G. *Phys. Rev.* **1929**, *34*, 553–573.

(11) Saue, T.; Fægri Fr., K.; Helgaker, T.; Gropen, O. *Mol. Phys.* **1997**, *91*, 937–950.

(12) Saue, T.; Jensen, H. J. Aa. *J. Chem. Phys.* **1999**, *111*, 6211–6222.

(13) Yanai, T.; Nakajima, T.; Ishikawa, Y.; Hirao, K. *J. Chem. Phys.* **2001**, *114*, 6526–6538.

(14) Kelley, M. S.; Shiozaki, T. *J. Chem. Phys.* **2013**, *138*, 204113.

(15) Liu, W.; Hong, G.; Dai, D.; Li, L.; Dolg, M. *Theor. Chem. Acc.* **1997**, *96*, 75–83.

(16) Yanai, T.; Iikura, H.; Nakajima, T.; Ishikawa, Y.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 8267–8273.

(17) Saue, T.; Helgaker, T. *J. Comput. Chem.* **2002**, *23*, 814–823.

(18) Quiney, H. M.; Belanzoni, P. *J. Chem. Phys.* **2002**, *117*, 5550–5563.

(19) Jensen, H. J. Aa.; Dyall, K. G.; Saue, T.; Fægri Jr., K. *J. Chem. Phys.* **1996**, *104*, 4083–4097.

(20) Thyssen, J.; Fleig, T.; Jensen, H. J. Aa. *J. Chem. Phys.* **2008**, *129*, 034109.

(21) Bates, J. E.; Shiozaki, T. *J. Chem. Phys.* **2015**, *142*, 044112.

(22) Laerdahl, K. J.; Saue, T.; Fægri Jr., K. *Theor. Chem. Acc.* **1997**, *97*, 177–184.

(23) Abe, M.; Yanai, T.; Nakajima, T.; Hirao, K. *Chem. Phys. Lett.* **2004**, *388*, 68–73.

(24) Fleig, T.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2003**, *119*, 2963–2971.

(25) Fleig, T.; Jensen, H. J. Aa.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2006**, *124*, 104106.

(26) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. *J. Chem. Phys.* **2010**, *132*, 014108.

(27) Visscher, L.; Dyall, K. G.; Lee, T. J. *Int. J. Quantum Chem.* **1995**, *56*, 411–419.

(28) Visscher, L.; Lee, T. J.; Dyall, K. G. *J. Chem. Phys.* **1996**, *105*, 8769–8776.

(29) Nataraj, H. S.; Kállay, M.; Visscher, L. *J. Chem. Phys.* **2010**, *133*, 234109.

(30) Cheng, L.; Gauss, J. *J. Chem. Phys.* **2011**, *135*, 084114.

(31) Dyall, K. G. *J. Chem. Phys.* **1997**, *106*, 9618–9626.

(32) Kutzelnigg, W.; Liu, W. *J. Chem. Phys.* **2005**, *123*, 241102.

(33) Iliaš, M.; Saue, T. *J. Chem. Phys.* **2007**, *126*, 064102.

(34) Cheng, L.; Gauss, J. *J. Chem. Phys.* **2011**, *135*, 084114.

(35) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89–155.

(36) Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742–3748.

(37) Chang, C.; Pelissier, M.; Durand, P. *Phys. Scr.* **1986**, *34*, 394.

(38) Rutkowski, A. *J. Phys. B* **1986**, *19*, 149.

(39) Stopkowicz, S.; Gauss, J. *J. Chem. Phys.* **2008**, *129*, 164119.

(40) Barysz, M.; Sadlej, A. J. *J. Chem. Phys.* **2002**, *116*, 2696–2704.

(41) Dyall, K. G. *J. Chem. Phys.* **1994**, *100*, 2118–2127.

(42) Visscher, L.; Saue, T. *J. Chem. Phys.* **2000**, *113*, 3996–4002.

(43) Fleig, T.; Visscher, L. *Chem. Phys.* **2005**, *311*, 113 – 120.

(44) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. *J. Chem. Phys.* **2008**, *128*, 014108.

(45) Cheng, L.; Stopkowicz, S.; Gauss, J. *J. Chem. Phys.* **2013**, *139*, 214114.

(46) Werner, H.-J. *Adv. Chem. Phys.* **2007**, *69*, 1–62.

(47) Shepard, R. *Adv. Chem. Phys.* **2007**, *69*, 63–200.

(48) Kim, Y. S.; Lee, Y. S. *J. Chem. Phys.* **2003**, *119*, 12169–12178.

(49) Shiozaki, T.; Mizukami, W. *J. Chem. Theory Comput.* **2015**, *11*, 4733–4739.

(50) Visscher, L.; Saue, T. *J. Chem. Phys.* **2000**, *113*, 3996–4002.

(51) Werner, H.-J.; Reinsch, E.-A. *J. Chem. Phys.* **1982**, *76*, 3144–3156.

(52) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(53) Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481 – 488.

(54) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. *Chem. Rev.* **2012**, *112*, 108–181.

(55) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. *WIREs Comput. Mol. Sci.* **2013**, *3*, 176–197.

(56) Hinze, J. *J. Chem. Phys.* **1973**, *59*, 6424–6432.

(57) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157–173.

(58) Eade, R. H. A.; Robb, M. A. *Chem. Phys. Lett.* **1981**, *83*, 362–368.

(59) Meier, U.; Staemmler, V. *Theor. Chem. Acc.* **1989**, *76*, 95–111.

(60) Frisch, M.; Ragazos, I. N.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Lett.* **1992**, *189*, 524–528.

(61) Levy, B.; Berthier, G. *Int. J. Quantum Chem.* **1968**, *2*, 307–319.

(62) Siegbahn, P. E. M.; Almlöf, J.; Heiberg, A.; Roos, B. O. *J. Chem. Phys.* **1981**, *74*, 2384–2396.

(63) Werner, H.-J.; Meyer, W. *J. Chem. Phys.* **1981**, *74*, 5794–5801.

(64) Jørgensen, P.; Swanstrøm, P.; Yeager, D. L. *J. Chem. Phys.* **1983**, *78*, 347–356.

(65) Jensen, H. J. Aa.; Jørgensen, P. *J. Chem. Phys.* **1984**, *80*, 1204–1214.

(66) Jensen, H. J. Aa.; Ågren, H. *Chem. Phys. Lett.* **1984**, *110*, 140–144.

(67) Jensen, H. J. Aa.; Ågren, H. *Chem. Phys.* **1986**, *104*, 229–250.

(68) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053–5063.

(69) Fletcher, R. *Practical Methods of Optimization, Second Edition*; Wiley, New York, 1987; Chapter 5.2, pp 100–107.

(70) Jensen, H. J. Aa.; Jørgensen, P.; Ågren, H. *J. Chem. Phys.* **1987**, *87*, 451–466.

(71) Visscher, L.; Aerts, P. J. C.; Visser, O.; Nieuwpoort, W. C. *Int. J. Quantum Chem.* **1991**, *40*, 131–139.

(72) Sucher, J. *Phys. Rev. A* **1980**, *22*, 348–362.

(73) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley, New York, 2000; Chapter 12.4, pp 616–619.

(74) Hylleraas, E. A.; Undheim, B. *Z. Phys.* **1930**, *65*, 759–772.

(75) MacDonald, J. K. L. *Phys. Rev.* **1933**, *43*, 830–833.

(76) CFOUR, a quantum chemical program package written by J. F. Stanton, J. Gauss, M. E. Harding, P. G. Szalay with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, L. Cheng, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, F. Lipparini, D. A. Matthews, T. Metzroth, L. A. Mück, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

(77) Handy, N. C. *Chem. Phys. Lett.* **1980**, *74*, 280–283.

(78) Knowles, P. J.; Handy, N. C. *Chem. Phys. Lett.* **1984**, *111*, 315–321.

(79) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. Aa. *J. Chem. Phys.* **1988**, *89*, 2185–2192.

(80) Bendazzoli, G. L.; Evangelisti, S. *J. Chem. Phys.* **1993**, *98*, 3141–3150.

(81) Gagliardi, L.; Bendazzoli, G. L.; Evangelisti, S. *J. Comput. Chem.* **1997**, *18*, 1329–1343.

(82) Brent, R. P. *Algorithms for Minimisation without Derivatives, Second Edition*; Dover Publications, Mineola, NY, 2002; Chapter 4, pp 47–56.

(83) Davidson, E. R. *J. Comput. Phys.* **1975**, *17*, 87–94.

(84) Liu, B. *Numerical Algorithms in Chemistry: Algebraic Methods, Proceedings from a workshop of the National Resource for Computational Chemistry*; Berkeley, 1978; pp 49–52.

(85) Olsen, J.; Jørgensen, P.; Simons, J. *Chem. Phys. Lett.* **1990**, *169*, 463–472.

(86) Gill, P. E.; Murray, W.; Saunders, M. A.; Wright, M. H. *Linear Algebra Appl.* **1987**, *88*, 239–270.

(87) Saunders, M. LUSOL: Sparse LU for Ax=b. The package is freely available at `https://github.com/nwh/lusol`. Last accessed February 15th 2016.

(88) Sleijpen, G. L. G.; der Vorst, H. A. V. *SIAM J. Matrix Anal. Appl.* **1996**, *17*, 401–425.

(89) Saad, Y.; Schultz, M. H. *SIAM J. Sci. Stat. Comp.* **1986**, *7*, 856–869.

(90) Dunning Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(91) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.

(92) Kim, J.; Kim, J.; Ihee, H. *J. Phys. Chem. A* **2013**, *117*, 3861–3868.

(93) Rives, A. B.; Fenske, R. F. *J. Chem. Phys.* **1981**, *75*, 1293–1302.

(94) Blomberg, M. R. A.; Brandemark, U. B.; Siegbahn, P. E. M.; Mathisen, K. B.; Karlström, G. *J. Phys. Chem.* **1985**, *89*, 2171–2180.

(95) Blomberg, M. R. A.; Siegbahn, P. E. M.; Lee, T. J.; Rendell, A. P.; Rice, J. E. *J. Chem. Phys.* **1991**, *95*, 5898–5905.

(96) Persson, B. J.; Roos, B. O.; Pierloot, K. *J. Chem. Phys.* **1994**, *101*, 6810–6821.

(97) Xu, X.; Lü, X.; Wang, N.; Zhang, Q.; Ehara, M.; Nakatsuji, H. *Int. J. Quantum Chem.* **1999**, *72*, 221–231.

(98) Yamazaki, E.; Okabayashi, T.; Tanimoto, M. *J. Am. Chem. Soc.* **2004**, *126*, 1028–1029.

(99) Filatov, M. *Chem. Phys. Lett.* **2003**, *373*, 131–135.

(100) Walker, N. R.; Hui, J. K.-H.; Gerry, M. C. L. *J. Phys. Chem. A* **2002**, *106*, 5803–5808.

(101) Tremblay, B.; Manceron, L. *Chem. Phys.* **1999**, *250*, 187–197.

(102) Basch, H.; Cohen, D. *J. Am. Chem. Soc.* **1983**, *105*, 3856–3860.

(103) Evans, C. J.; Gerry, M. C. L. *J. Phys. Chem. A*. **2001**, *105*, 9659–9663.

(104) Okabayashi, T.; Yamamoto, T.; Okabayashi, E. Y.; Tanimoto, M. *J. Phys. Chem. A* **2011**, *115*, 1869–1877.

(105) Zhou, M.; Andrews, L.; Bauschlicher Jr., C. W. *Chem. Rev.* **2001**, *101*, 1931–1962, and references therein.

(106) McKinlay, R. G.; Almeida, N. M. S.; Coe, J. P.; Paterson, M. J. *J. Phys. Chem. A* **2015**, *119*, 10076–10083.

(107) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. *J. Phys. Chem. A* **2005**, *109*, 6575–6579.

Figure 5: Table of contents and abstract graphics