Multiconfigurational short-range density-functional theory for open-shell systems

Erik Donovan Hedegård,[1,a)] Julien Toulouse,[2] and Hans Jørgen Aagaard Jensen[3,b)]
[1]*Department of Theoretical Chemistry, Lund University, Kemicentrum P.O. Box 124, SE-221 00 Lund, Sweden*
[2]*Laboratoire de Chimie Théorique, Sorbonne Université and CNRS, Paris, France*
[3]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense M, Denmark*

(Received 13 November 2017; accepted 14 May 2018; published online 5 June 2018)

Many chemical systems cannot be described by quantum chemistry methods based on a single-reference wave function. Accurate predictions of energetic and spectroscopic properties require a delicate balance between describing the most important configurations (static correlation) and obtaining dynamical correlation efficiently. The former is most naturally done through a multiconfigurational (MC) wave function, whereas the latter can be done by, e.g., perturbation theory. We have employed a different strategy, namely, a hybrid between multiconfigurational wave functions and density-functional theory (DFT) based on range separation. The method is denoted by MC short-range DFT (MC–srDFT) and is more efficient than perturbative approaches as it capitalizes on the efficient treatment of the (short-range) dynamical correlation by DFT approximations. In turn, the method also improves DFT with standard approximations through the ability of multiconfigurational wave functions to recover large parts of the static correlation. Until now, our implementation was restricted to closed-shell systems, and to lift this restriction, we present here the generalization of MC–srDFT to open-shell cases. The additional terms required to treat open-shell systems are derived and implemented in the DALTON program. This new method for open-shell systems is illustrated on dioxygen and $[Fe(H_2O)_6]^{3+}$. *Published by AIP Publishing.* https://doi.org/10.1063/1.5013306

I. INTRODUCTION

Quantum chemistry (QC) methods have become paramount to gain insight in experimental investigations and to support findings from experimental studies. Moreover, QC methods can be employed to predict novel molecular systems with a given desired property. Naturally, the above uses of theoretical methods require that these methods are both efficient and accurate, also in cases where reference data might not exist or be inconclusive. Density-functional theory (DFT) has to a large extent fulfilled these requirements[1] and can for many systems be employed routinely. Still, DFT with standard approximations is known to fail for systems that cannot be described by a single-reference wave function.[2,3] Such systems are often characterized by dense orbital manifolds and several (near)-degenerate states. In wave-function theory (WFT), such systems are best described with a multiconfigurational wave function.[4] These exist in a number of different forms of which many have in common the definition of a *complete active space* CAS(m, n) of m electrons in n orbitals. In this active space, all configurations (fulfilling additional spin and symmetry constraints) are included. This is usually combined with optimization of orbital parameters in what is denoted by a complete-active-space self-consistent field (CASSCF) procedure. Unfortunately, the computational demands increase dramatically with the size of the active space, and many systems require active spaces beyond the current limitations to give physically meaningful results. In recent years, several groups have focused on lifting the limitations for the size of active space with methods such as the density-matrix renormalization group (DMRG),[5–8] quantum Monte Carlo (QMC),[9–11] restricted-active-space (RAS),[12] or generalized-active-space (GAS) methods.[13–15] Yet, even with extended active orbital spaces, essential parts of the remaining dynamical electron correlation cannot be obtained, except for the smallest systems where the number of occupied valence orbitals is small. Typically, dynamical correlation for such multireference systems is obtained *after* initially obtaining a correct representation of a zeroth-order Hamiltonian (including static correlation). The exact nature of the subsequent steps responsible for recovering dynamical correlation depends on the chosen method, but well-known examples are multireference perturbation theories[16–20] such as complete-active-space second-order perturbation theory (CASPT2)[16,17] and n-electron valence state perturbation theory (NEVPT2).[20]

We have in a number of recent papers employed a different strategy, namely, a hybrid between MCSCF and DFT that capitalizes on the fairly efficient treatment of the (short-range) dynamical correlation by semi-local DFT approximations and the ability of CASSCF or more general MCSCF models to recover large parts of the static correlation. A number of different CAS-DFT hybrid methods have been suggested,[21–30] many of which are under active development. There is one important difference between these methods: some add DFT after initial optimization of the multiconfigurational wave function,[23,24]

[a)]Electronic mail: erik.hedegard@teokem.lu.se
[b)]Electronic mail: hjj@sdu.dk

whereas others optimize the DFT and wave function parts simultaneously.[21,22,25–30] The method reported here belongs to the latter kind. A challenge for all CAS-DFT hybrids is to avoid double counting of electron correlation, as the CAS wave functions will invariably include some of the dynamical correlation. The method that we discuss in this paper, and that we have been developing since several years,[27,29–36] stringently avoids this double counting by means of a range separation of the two-electron repulsion operator.[25,37] The *long-range* part is then described by WFT, whereas the *short-range* part is described by a tailored short-range DFT functional.[38] When employing an MC(SCF) wave function as the long-range component, we have dubbed the method multiconfigurational short-range DFT (MC–srDFT). For the common choice of CAS for the MCSCF part, this becomes CAS–srDFT. However, it should be emphasized that the method of range separation also allows for other wave-function *ansätze*: CI–srDFT,[26,39] MP2–srDFT,[40,41] RPA-srDFT,[42,43] CC–srDFT,[44] RAS–srDFT, NEVPT2–srDFT,[45] and DMRG–srDFT[32] methods have all been implemented.

Our MC–srDFT implementation in a development version of DALTON[46,47] employs a general implementation for the long-range MCSCF wave function that allows any spin multiplicity. However, the implementation has until now been restricted to short-range functionals depending only on the total density and not on the spin density. Calculations have accordingly been done on closed-shell systems only. It should in this regard be noted that introducing a dependence on the spin density in the DFT part is not necessary from a theoretical point of view where the energy is solely defined from the total density. Yet, approximate semi-local functionals must depend on the spin density to be accurate for open-shell systems, and a semi-local srDFT implementation based solely on the total density in practice excludes useful calculations on many open-shell target systems where we expect MC–srDFT to be able to perform well. For example, many transition metals form open-shell complexes, and these are notorious for displaying multireference character.

Since the original formulation of MC–srDFT,[39] several srDFT functionals have been generalized to include the spin density.[48,49] In this paper, we extend the MC–srDFT method with these spin-dependent functionals, both for a local-density approximation (LDA) and a generalized-gradient approximation (GGA). This enables meaningful calculations for open-shell systems, which in particular will be beneficial for our endeavors to describe transition-metal complexes with CAS–srDFT or GAS–srDFT as a computationally cheaper alternative to CASPT2, RASPT2, and NEVPT2.

As a first test of our novel implementation, we investigate two prototypical open-shell molecules: the O_2 molecule and the transition metal complex $[Fe(H_2O)_6]^{3+}$. We focus here on rather simple systems which allow for a thorough investigation and where literature data are available. For both the $^3\Sigma_g^-$ ground-state of O_2 and the spin-state splitting in $[Fe(H_2O)_6]^{3+}$, we also carry out a small study on the dependence of CAS–srDFT on the range-separation parameter μ, similar to what has previously been done for a number of small, closed-shell p-block molecules and atoms.[27] It should be emphasized that although O_2 is a prototypical open-shell system for which a wealth of literature data is available,[50] the spin-state splitting between the $^3\Sigma_g^-$ ground-state and the first singlet state, $^1\Delta_g$, is surprisingly demanding in terms of the theoretical treatment.[51–53] Likewise, the spin-state splitting between the lowest electronic states, ^6A_g and $^4T_{1g}$ (assuming idealized, octahedral symmetry) in $[Fe(H_2O)_6]^{3+}$, has turned out to be challenging; there is currently no consensus on which method is the most accurate, as both CASPT2 and CCSD(T) have been shown to overestimate the spin-state splitting somewhat.[54] Accordingly, both O_2 and $[Fe(H_2O)_6]^{3+}$ have low-lying excited states with different spin multiplicity than the ground-state, and for both a high-level method is evidently needed to obtain correct results.

The paper is organized as follows. In Sec. II, we describe the necessary theory for the extension of MC–srDFT to srDFT functionals that include the spin density as well as the gradient of the spin density, and in Sec. III we describe implementation details. The theory and implementation details for the closed-shell singlet case are obviously a special case, which previously only has been published in the Ph.D. thesis by Pedersen.[39] Next, we provide the computational details that were employed for the test calculations (Sec. IV), and in Sec. V we discuss the results. Finally, we give a brief conclusion and outlook in Sec. VI.

II. RANGE-SEPARATED MULTICONFIGURATIONAL DENSITY FUNCTIONAL THEORY

The MC–srDFT method employs a range separation of the two-electron operator[25,37]

$$\hat{g}(1,2) = \hat{g}^{\mathrm{lr}}(1,2) + \hat{g}^{\mathrm{sr}}(1,2), \tag{1}$$

in which the exact definition of $\hat{g}^{\mathrm{lr}}(1,2)$ and $\hat{g}^{\mathrm{sr}}(1,2)$ can differ.[55] In this work, we exclusively use the error function for the range separation

$$\hat{g}^{\mathrm{lr}}(1,2) = \frac{\mathrm{erf}(\mu r_{12})}{r_{12}} \quad \text{and} \quad \hat{g}^{\mathrm{sr}}(1,2) = \frac{1-\mathrm{erf}(\mu r_{12})}{r_{12}}, \tag{2}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ and μ is the range-separation parameter, given in reciprocal bohr. Equation (1) is given in atomic units which we will employ throughout this paper. The energy of a range-separated WFT-DFT hybrid is then given by

$$\begin{aligned} E(\boldsymbol{\lambda}) &= E_{\mathrm{lr}}(\boldsymbol{\lambda}) + E_{\text{sr-H}}[\rho_C] + E_{\text{sr-xc}}[\rho_C, \rho_S] \\ &= \langle \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) | \hat{H}^{\mathrm{lr}} | \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) \rangle + E_{\text{sr-H}}[\rho_C] + E_{\text{sr-xc}}[\rho_C, \rho_S], \end{aligned} \tag{3}$$

where $\boldsymbol{\lambda}$ are the wave-function parameters, to be defined below. We denote the electron charge density as ρ_C and the spin density as ρ_S, and in order not to overload the notation we let it be implicit that they depend on the wave-function parameters $\boldsymbol{\lambda}$: $\rho_C \equiv \rho_C(\mathbf{r}, \boldsymbol{\lambda})$ and $\rho_S \equiv \rho_S(\mathbf{r}, \boldsymbol{\lambda})$. We define the short-range Hartree and exchange-correlation functionals, $E_{\text{sr-H}}[\rho_C]$ and $E_{\text{sr-xc}}[\rho_C, \rho_S]$, below after first defining the quantities involved in the long-range part of the energy in Eq. (3). The long-range Hamiltonian $\hat{H}^{\mathrm{lr}}$ takes a form similar to the regular electronic (non- or scalar-relativistic) Hamiltonian in second quantization,

$$\hat{H}^{\mathrm{lr}} = \sum_{pq} h_{pq} \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs} g^{\mathrm{lr}}_{pq,rs} \hat{e}_{pq,rs} + V_{\mathrm{nn}}, \tag{4}$$

with the standard one-electron integrals h_{pq} and nuclei-nuclei interaction energy V_{nn}, but with the two-electron Coulomb integrals replaced by integrals with the modified long-range interaction in Eq. (2),

$$g^{\mathrm{lr}}_{pq,rs} = \langle \phi_p \phi_r | \hat{g}^{\mathrm{lr}}(1,2) | \phi_q \phi_s \rangle, \tag{5}$$

in an orthonormal basis of spatial orbitals $\{\phi_p\}$. The second-quantized operators in Eq. (4) have their usual meaning: $\hat{E}_{pq} = \hat{a}^\dagger_{p\alpha}\hat{a}_{q\alpha} + \hat{a}^\dagger_{p\beta}\hat{a}_{q\beta}$ is the singlet excitation operator and $\hat{e}_{pq,rs} = \hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}$. The indices p, q, r, s denote general orbitals. For the MC(SCF) wave-function parameters, we use the notation $\boldsymbol{\lambda}^T = (\boldsymbol{d}^T, \boldsymbol{\kappa}^T)$, where the row vectors $\boldsymbol{d}^T = \{d_j\}$ and $\boldsymbol{\kappa}^T = \{\kappa_{pq}\}$ designate the configuration coefficients and the orbital rotation parameters, respectively. The real-valued MC wave function is parameterized as

$$|\Psi^{\mathrm{lr}}(\boldsymbol{\lambda})\rangle = e^{-\hat{\kappa}} \left(\frac{|0\rangle + \hat{\mathcal{P}}|\,\boldsymbol{d}\,\rangle}{\sqrt{1 + \langle\,\boldsymbol{d}\,|\hat{\mathcal{P}}|\,\boldsymbol{d}\,\rangle}} \right), \tag{6}$$

where $|0\rangle$ denotes a normalized reference state

$$|0\rangle = \sum_j c_j |j\rangle, \tag{7}$$

while

$$|\boldsymbol{d}\rangle = \sum_j d_j |j\rangle \tag{8}$$

is a configuration correction and $\hat{\mathcal{P}} = 1 - |0\rangle\langle 0|$ is the projection operator onto the complement to the reference state $|0\rangle$. The $\hat{\kappa}$ operator in Eq. (6) is the usual antisymmetric real singlet orbital-rotation operator

$$\hat{\kappa} = \sum_{pq} \kappa_{pq} \hat{E}_{pq} = \sum_{p>q} \kappa_{pq} \left(\hat{E}_{pq} - \hat{E}_{qp} \right) \equiv \sum_{p>q} \kappa_{pq} \hat{E}^-_{pq}. \tag{9}$$

The charge-density- and spin-density-dependent terms in Eq. (3) can now be expressed in a second-quantization formulation[56,57] in terms of their associated density operators

$$\hat{\rho}_X(\mathbf{r}) = \sum_{pq} \Omega_{pq}(\mathbf{r}) \hat{O}^X_{pq}, \tag{10}$$

with $\Omega_{pq}(\mathbf{r}) = \phi^*_p(\mathbf{r})\phi_q(\mathbf{r})$. We have in Eq. (10) introduced a nomenclature that will turn convenient in Secs. II A and II B for equations which are otherwise identical for charge-density and spin-density operators. In this nomenclature, $X = C$ is used for the regular charge-density operator ($\hat{O}^C_{pq} \equiv \hat{E}_{pq}$), while $X = S$ denotes the spin-density operator

$$\hat{O}^S_{pq} \equiv \hat{T}_{pq} = \hat{a}^\dagger_{p\alpha}\hat{a}_{q\alpha} - \hat{a}^\dagger_{p\beta}\hat{a}_{q\beta}, \tag{11}$$

which is required to describe spin-density effects (in a non-relativistic framework). We note that for spin-restricted models as considered here, the charge-density operator is a singlet operator and the spin-density operator is a triplet operator of $M_S = 0$ type. The wave function can be of any spin symmetry S; however, as in Kohn-Sham DFT, in practice, the wave function should correspond to a spin component $M_S = S$ or $M_S = -S$ to give an appropriate spin density for the approximate spin-polarized exchange-correlation functional. The electron charge density $\rho_C(\mathbf{r}, \boldsymbol{\lambda})$ and electron spin density $\rho_S(\mathbf{r}, \boldsymbol{\lambda})$ are obtained as expectation values according to

$$\rho_X(\mathbf{r}, \boldsymbol{\lambda}) = \langle \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) | \hat{\rho}_X(\mathbf{r}) | \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) \rangle = \sum_{pq} \Omega_{pq}(\mathbf{r}) D^X_{pq}(\boldsymbol{\lambda}), \tag{12}$$

where $D^X_{pq}(\boldsymbol{\lambda})$ is the (p, q) element of the one-electron reduced (spin-)density matrix

$$D^X_{pq}(\boldsymbol{\lambda}) = \langle \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) | \hat{O}^X_{pq} | \Psi^{\mathrm{lr}}(\boldsymbol{\lambda}) \rangle. \tag{13}$$

The $E_{\text{sr-H}}[\rho_C]$ and $E_{\text{sr-xc}}[\rho_C, \rho_S]$ terms in Eq. (3) can now be defined. The former is the short-range Hartree energy which depends only on the total density matrix

$$E_{\text{sr-H}}[\rho_C] = \frac{1}{2} \sum_{pq,rs} D^C_{pq}(\boldsymbol{\lambda})\, g^{\mathrm{sr}}_{pqrs}\, D^C_{rs}(\boldsymbol{\lambda}) \equiv \frac{1}{2} \sum_{pq} D^C_{pq}(\boldsymbol{\lambda}) j^{\mathrm{sr}}_{pq}(\boldsymbol{\lambda}), \tag{14}$$

where g^{sr}_{pqrs} are the short-range two-electron integrals [defined with the short-range interaction in Eq. (2)]. The final term, $E_{\text{sr-xc}}[\rho_C, \rho_S]$, in Eq. (3) is the short-range exchange-correlation functional. This term has an explicit dependence on both the charge and spin densities

$$E_{\text{sr-xc}}[\rho_C, \rho_S] = \int e_{\text{sr-xc}}(\rho_C(\mathbf{r}, \boldsymbol{\lambda}), \rho_S(\mathbf{r}, \boldsymbol{\lambda}))\mathrm{d}\mathbf{r}, \tag{15}$$

in accordance with Ref. 56, where $e_{\text{sr-xc}}(\rho_C(\mathbf{r}, \boldsymbol{\lambda}), \rho_S(\mathbf{r}, \boldsymbol{\lambda}))$ is the short-range exchange-correlation energy density. It should be noted that the expression in Eq. (15) assumes the local-density approximation of the short-range exchange-correlation functional (srLDA).[48] For brevity we only consider srLDA here, but we have also implemented the code for functionals based on a generalized-gradient approximation (GGA). The spin-dependent GGAs additionally depend on the electron charge density gradient and the electron spin density gradient. The additional terms are structurally similar to the srLDA terms and will be given in the Appendix.

The MC–srDFT wave function is optimized with the restricted-step second-order MCSCF optimization algorithm[58–62] as implemented in DALTON.[46,47] This algorithm is based on a second-order Taylor expansion of the electronic energy in the wave-function parameters, $\boldsymbol{\lambda}$, around $\boldsymbol{\lambda} = 0$,

$$E(\boldsymbol{\lambda}) = E_0 + \boldsymbol{g}^T\boldsymbol{\lambda} + \frac{1}{2}\boldsymbol{\lambda}^T\boldsymbol{H}\boldsymbol{\lambda} + \cdots. \tag{16}$$

The electronic gradient, $\boldsymbol{g}$, and electronic Hessian, $\boldsymbol{H}$, are blocked according to the configurational and orbital parameters of the wave function. Thus the gradient reads

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g}^c \\ \boldsymbol{g}^o \end{pmatrix}, \tag{17}$$

and each gradient element (configurational or orbital) has both WFT (lr) and DFT (sr) contributions,

$$\begin{aligned} g_i &= g_{\text{lr},i} + g_{\text{sr-H},i} + g_{\text{sr-xc},i} \\ &= \frac{\partial E_{\text{lr}}}{\partial \lambda_i} + \frac{\partial E_{\text{sr-H}}[\rho_C]}{\partial \lambda_i} + \frac{\partial E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \lambda_i}. \end{aligned} \tag{18}$$

The electronic Hessian in Eq. (16) is evaluated as $\boldsymbol{H} = \boldsymbol{PKP}$, where $\boldsymbol{P}$ denotes the matrix representation of the $\hat{\mathcal{P}}$ operator in Eq. (6) and $\boldsymbol{K}$ has WFT and DFT contributions (see Ref. 62)

$$\begin{aligned} K_{ij} &= K_{\text{lr},ij} + K_{\text{sr-H},ij} + K_{\text{sr-xc},ij} \\ &= \frac{\partial^2 E_{\text{lr}}}{\partial \lambda_i \partial \lambda_j} + \frac{\partial^2 E_{\text{sr-H}}[\rho_C]}{\partial \lambda_i \partial \lambda_j} + \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \lambda_i \partial \lambda_j}. \end{aligned} \tag{19}$$

In the actual implementation, the Hessian matrix is never constructed explicitly. Instead, Hessian contributions to the wave-function optimization process are obtained in a direct fashion based on trial vectors

$$\boldsymbol{\sigma}_n = \boldsymbol{PKP}\,\boldsymbol{b}_n = \boldsymbol{PK}\,\boldsymbol{b}_n = \boldsymbol{P}\begin{pmatrix} \boldsymbol{K}^{cc} & \boldsymbol{K}^{co} \\ \boldsymbol{K}^{oc} & \boldsymbol{K}^{oo} \end{pmatrix}\begin{pmatrix} \boldsymbol{b}_n^c \\ \boldsymbol{b}_n^o \end{pmatrix}, \tag{20}$$

where the individual contributions to $\boldsymbol{\sigma}_n$ can be written in terms of modified Fock matrices (see Sec. III) and the second equal sign stems from that we require each trial vector to fulfill $\boldsymbol{P}\,\boldsymbol{b}_n = \boldsymbol{b}_n$. The explicit expressions for the new short-range DFT contributions to the individual gradient and $\boldsymbol{\sigma}_n$ element types are derived below. The MCSCF long-range contributions are equivalent to the terms in the regular MCSCF method which together with the second-order optimization algorithm are documented in the original studies describing the DALTON implementation.[58–62]

A. The sr-DFT contributions to the electronic gradient

The short-range exchange-correlation contributions to the gradient in Eq. (18) can be obtained as

$$g_{\text{sr-xc},i} = \frac{\partial E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \lambda_i} = \int \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \lambda_i} \mathrm{d}\mathbf{r}. \tag{21}$$

The derivative within the kernel in Eq. (21) is obtained through the chain rule

$$\begin{aligned} \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \lambda_i} &= \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_C} \frac{\partial \rho_C(\mathbf{r})}{\partial \lambda_i} + \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_S} \frac{\partial \rho_S(\mathbf{r})}{\partial \lambda_i} \\ &= \sum_{pq} \left[\left(\frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_C} \Omega_{pq}(\mathbf{r}) \right) \frac{\partial D^C_{pq}(\boldsymbol{\lambda})}{\partial \lambda_i} + \left(\frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_S} \Omega_{pq}(\mathbf{r}) \right) \frac{\partial D^S_{pq}(\boldsymbol{\lambda})}{\partial \lambda_i} \right], \end{aligned} \tag{22}$$

where $\rho_C(\mathbf{r})$ and $\rho_S(\mathbf{r})$ from Eq. (12) have been inserted. As usual in MCSCF schemes, the orbital and configurational parts of the gradient leads to different computational expressions. The orbital part of the short-range exchange-correlation gradient becomes

$$\frac{\partial E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \kappa_{rs}} = g^C_{\text{sr-xc},rs} + g^S_{\text{sr-xc},rs}, \tag{23}$$

where the two gradient terms

$$g^C_{\text{sr-xc},rs} = \langle 0 | \left[\hat{E}^-_{rs}, \hat{V}^{C,g}_{\text{sr-xc}} \right] | 0 \rangle = 2 \langle 0 | \left[\hat{E}_{rs}, \hat{V}^{C,g}_{\text{sr-xc}} \right] | 0 \rangle, \tag{24a}$$

$$g^S_{\text{sr-xc},rs} = \langle 0 | \left[\hat{E}^-_{rs}, \hat{V}^{S,g}_{\text{sr-xc}} \right] | 0 \rangle = 2 \langle 0 | \left[\hat{E}_{rs}, \hat{V}^{S,g}_{\text{sr-xc}} \right] | 0 \rangle, \tag{24b}$$

are both defined in terms of an effective operator ($X = C$ and $X = S$, respectively)

$$\begin{aligned} \hat{V}^{X,g}_{\text{sr-xc}} &= \sum_{pq} \left(\int \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_X} \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \hat{O}^X_{pq} \\ &\equiv \sum_{pq} V^{X,g}_{\text{sr-xc},pq} \hat{O}^X_{pq}. \end{aligned} \tag{25}$$

We note that the $g^C_{\text{sr-xc},rs}$ term in Eq. (24a) is the term already present in the closed-shell formalism, although in that case the functional kernel in the effective operator does not depend on ρ_S. The $g^S_{\text{sr-xc},rs}$ term in Eq. (24b) exclusively occurs in the open-shell formalism. The two gradient terms are commonly denoted by $\mathbf{g}^X_{\text{sr-xc}}$. The configurational part of the gradient can be defined similarly as

$$\frac{\partial E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial d_j} = g^C_{\text{sr-xc},j} + g^S_{\text{sr-xc},j}, \tag{26}$$

where

$$g^X_{\text{sr-xc},j} = 2\left(\langle j | \hat{V}^{X,g}_{\text{sr-xc}} | 0 \rangle - c_j \langle 0 | \hat{V}^{X,g}_{\text{sr-xc}} | 0 \rangle \right), \tag{27}$$

and we again have employed the effective operators defined in Eq. (25).

Finally, we also briefly describe the gradient contribution from the short-range Hartree term, $E_{\text{sr-H}}[\rho_C]$, in Eq. (14),

$$g_{\text{sr-H},i} = \frac{\partial E_{\text{sr-H}}[\rho_C]}{\partial \lambda_i} = \sum_{pq} j^{\text{sr}}_{pq}(\boldsymbol{\lambda}) \frac{\partial D^C_{pq}(\boldsymbol{\lambda})}{\partial \lambda_i}. \tag{28}$$

The configurational and orbital gradient contributions are identical to Eqs. (24a) and (27), except that the effective operator

$$\hat{V}^g_{\text{sr-H}} = \sum_{pq} j^{\text{sr}}_{pq} \hat{E}_{pq} \tag{29}$$

replaces $\hat{V}^{C,g}_{\text{sr-xc}}$. Combined, we obtain the effective operators for the total short-range Hartree-exchange-correlation functional as

$$\hat{V}^{X,g}_{\text{sr-Hxc}} = \begin{cases} \hat{V}^g_{\text{sr-H}} + \hat{V}^{C,g}_{\text{sr-xc}}, & X = C \\ \hat{V}^{S,g}_{\text{sr-xc}}, & X = S \end{cases}. \tag{30}$$

The total gradient vector (from the short-range DFT part) can thus be obtained from

$$\mathbf{g}_{\text{sr-Hxc}} = \mathbf{g}^{C}_{\text{sr-Hxc}} + \mathbf{g}^{S}_{\text{sr-xc}}. \tag{31}$$

B. The sr-DFT contributions to the electronic Hessian sigma vectors

As for the gradient in Subsection II A, we focus on the exchange-correlation term since the second derivative of E_{lr} in Eq. (3) is equivalent to the terms from regular MCSCF,[58–61] and the Hartree term can be obtained from the total density matrix and has no explicit spin-dependent terms. The Hessian contributions from the exchange-correlation functional are given by

$$K_{\text{sr-xc},ij} = \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial\lambda_i \partial\lambda_j} = \int \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\lambda_i \partial\lambda_j} \mathrm{d}\mathbf{r}, \tag{32}$$

and from the chain rule

$$\begin{aligned}
\frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\lambda_i \partial\lambda_j} &= \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_C^2} \frac{\partial\rho_C(\mathbf{r})}{\partial\lambda_i} \frac{\partial\rho_C(\mathbf{r})}{\partial\lambda_j} + \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_C \partial\rho_S} \frac{\partial\rho_S(\mathbf{r})}{\partial\lambda_i} \frac{\partial\rho_C(\mathbf{r})}{\partial\lambda_j} \\
&+ \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_S \partial\rho_C} \frac{\partial\rho_C(\mathbf{r})}{\partial\lambda_i} \frac{\partial\rho_S(\mathbf{r})}{\partial\lambda_j} + \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_S^2} \frac{\partial\rho_S(\mathbf{r})}{\partial\lambda_i} \frac{\partial\rho_S(\mathbf{r})}{\partial\lambda_j} \\
&+ \frac{\partial e_{\text{xc}}\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_C} \frac{\partial^2\rho_C(\mathbf{r})}{\partial\lambda_i \partial\lambda_j} + \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_S} \frac{\partial^2\rho_S(\mathbf{r})}{\partial\lambda_i \partial\lambda_j}.
\end{aligned} \tag{33}$$

In practice, we utilize a direct Hessian technique [cf. Eq. (20)] where the quantity in Eq. (32) is contracted with configurational (b^c_j) or orbital (b^o_{pq}) trial vectors. The direct Hessian contributions from the $E_{\text{sr-xc}}[\rho_C, \rho_S]$ term thus consist of four different types,

$$\sum_j \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial d_i \partial d_j} b^c_j = \sigma^{C;c}_{\text{sr-xc},i} + \sigma^{S;c}_{\text{sr-xc},i}, \tag{34a}$$

$$\sum_j \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \kappa_{pq} \partial d_j} b^c_j = \sigma^{C;c}_{\text{sr-xc},pq} + \sigma^{S;c}_{\text{sr-xc},pq}, \tag{34b}$$

$$\sum_{r>s} \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial d_i \partial \kappa_{rs}} b^o_{rs} = \sigma^{C;o}_{\text{sr-xc},i} + \sigma^{S;o}_{\text{sr-xc},i}, \tag{34c}$$

$$\sum_{r>s} \frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial \kappa_{pq} \partial \kappa_{rs}} b^o_{rs} = \sigma^{C;o}_{\text{sr-xc},pq} + \sigma^{S;o}_{\text{sr-xc},pq}. \tag{34d}$$

The expressions in Eqs. (34a)–(34d) are obtained by inserting $\rho_C(\mathbf{r})$ and $\rho_S(\mathbf{r})$ from Eq. (12) into Eq. (33) and inserting the resulting expression into Eq. (32) and contracting with configurational or orbital trial vectors, respectively. We sketch the two steps below; the insertion of the densities gives

$$\begin{aligned}
\frac{\partial^2 E_{\text{sr-xc}}[\rho_C, \rho_S]}{\partial\lambda_i \partial\lambda_j} &= \sum_{X,Y=C,S} \sum_{pq,rs} \frac{\partial D^Y_{rs}(\boldsymbol{\lambda})}{\partial\lambda_j} \left(\int \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_Y \partial\rho_X} \Omega_{pq}(\mathbf{r}) \Omega_{rs}(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \frac{\partial D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i} \\
&+ \sum_{X=C,S} \sum_{pq} \left(\int \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_X} \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right) \frac{\partial^2 D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i \partial\lambda_j}.
\end{aligned} \tag{35}$$

By contraction of the above equation with the trial vectors, the configurational part becomes

$$\begin{aligned}
\sigma^{X;c}_{\text{sr-xc},i} &= \sum_{Y=C,S} \sum_{pq,rs} \left\{ \int \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_Y \partial\rho_X} \sum_j \left(\frac{\partial D^Y_{rs}(\boldsymbol{\lambda})}{\partial d_j} b^c_j \, \Omega_{rs}(\mathbf{r}) \right) \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right\} \frac{\partial D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i} \\
&+ \sum_{pq} \left\{ \int \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_X} \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right\} \sum_j \frac{\partial^2 D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i \partial d_j} b^c_j,
\end{aligned} \tag{36}$$

while the orbital part becomes

$$\begin{aligned}
\sigma^{X;o}_{\text{sr-xc},i} &= \sum_{Y=C,S} \sum_{pq,rs} \left\{ \int \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_Y \partial\rho_X} \sum_{t>u} \left(\frac{\partial D^Y_{rs}(\boldsymbol{\lambda})}{\partial \kappa_{tu}} b^o_{tu} \, \Omega_{rs}(\mathbf{r}) \right) \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right\} \frac{\partial D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i} \\
&+ \sum_{pq} \left\{ \int \frac{\partial e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial\rho_X} \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r} \right\} \sum_{t>u} \frac{\partial^2 D^X_{pq}(\boldsymbol{\lambda})}{\partial\lambda_i \partial \kappa_{tu}} b^o_{tu}.
\end{aligned} \tag{37}$$

The first term in both Eqs. (36) and (37) occurs due to the non-linearity of the srDFT energy functional. These terms can be formulated as transformed gradient terms with effective operators as summarized below. The last term in both Eqs. (36) and (37) is similar to the one-electron Hessian part for a regular MCSCF calculation, the only difference being that they employ different integrals (over the first-order derivative of the srDFT functional). To define the effective operators employed in the first term of the two equations, we define the transition density matrices for configurational trial vectors

$$D^{Y(1c)}_{rs} \equiv \sum_j \frac{\partial D^Y_{rs}(\boldsymbol{\lambda})}{\partial d_j} b^c_j = \sum_j \left(b^c_j \langle j|\hat{O}^Y_{rs}|0\rangle + \langle 0|\hat{O}^Y_{rs}|j\rangle b^c_j \right)$$
$$= \langle B|\hat{O}^Y_{rs}|0\rangle + \langle 0|\hat{O}^Y_{rs}|B\rangle, \tag{38}$$

where $|B\rangle$ is the state vector generated by the current trial vector: $|B\rangle = \sum_j b^c_j |j\rangle$. For the orbital trial vectors, we define the one-index transformed density matrix

$$D^{Y(1o)}_{rs} \equiv \sum_{t>u} \frac{\partial D^Y_{rs}(\boldsymbol{\lambda})}{\partial \kappa_{tu}} b^o_{tu} = \sum_t \left(D^Y_{ts} b^o_{rt} + D^Y_{rt} b^o_{st} \right). \tag{39}$$

The transformed configurational and orbital (spin-)density matrices, $\mathbf{D}^{Y(1c)}$ and $\mathbf{D}^{Y(1o)}$, are collectively denoted by $\mathbf{D}^{Y(1\lambda)}$. We can now define the integrals that comprise the effective operators in terms of the linearly transformed densities,

$$V^{X[1\lambda]}_{\text{sr-xc},pq}[\mathbf{D}^{Y(1\lambda)}] = \int \frac{\partial^2 e_{\text{sr-xc}}(\rho_C(\mathbf{r}), \rho_S(\mathbf{r}))}{\partial \rho_Y \partial \rho_X} \times \left(\sum_{rs} D^{Y(1\lambda)}_{rs} \Omega_{rs}(\mathbf{r}) \right) \Omega_{pq}(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{40}$$

which is a generic term for the integrals in the curvy brackets of the first term in Eqs. (36) and (37). With this generic definition of the integrals, we can define the two effective operators of either singlet or triplet types,

$$\hat{V}^{C[1\lambda]}_{\text{sr-xc}} = \sum_{pq} \left(V^{C[1\lambda]}_{\text{sr-xc},pq}[\mathbf{D}^{C(1\lambda)}] + V^{C[1\lambda]}_{\text{sr-xc},pq}[\mathbf{D}^{S(1\lambda)}] \right) \hat{E}_{pq}, \tag{41a}$$

$$\hat{V}^{S[1\lambda]}_{\text{sr-xc}} = \sum_{pq} \left(V^{S[1\lambda]}_{\text{sr-xc},pq}[\mathbf{D}^{C(1\lambda)}] + V^{S[1\lambda]}_{\text{sr-xc},pq}[\mathbf{D}^{S(1\lambda)}] \right) \hat{T}_{pq}. \tag{41b}$$

The direct Hessian terms in Eqs. (34a)–(34d) can now be written in the generic form as

$$\sigma^{X;c}_{\text{sr-xc},i} = 2\left(\langle i|\hat{V}^{X[1c]}_{\text{sr-xc}}|0\rangle - c_i \langle 0|\hat{V}^{X[1c]}_{\text{sr-xc}}|0\rangle \right) + 2\left(\langle i|\hat{V}^{X,g}_{\text{sr-xc}}|B\rangle - \langle 0|\hat{V}^{X,g}_{\text{sr-xc}}|0\rangle b^c_j \right), \tag{42a}$$

$$\sigma^{X;c}_{\text{sr-xc},pq} = 2\langle 0|\,[\hat{E}_{pq}, \hat{V}^{X[1c]}_{\text{sr-xc}}]\,|0\rangle + 2\langle 0|[\hat{E}^-_{pq}, \hat{V}^{X,g}_{\text{sr-xc}}]|B\rangle, \tag{42b}$$

$$\sigma^{X;o}_{\text{sr-xc},i} = 2\left(\langle i|\hat{V}^{X[1o]}_{\text{sr-xc}}|0\rangle - c_i \langle 0|\hat{V}^{X[1o]}_{\text{sr-xc}}|0\rangle \right) + 2\langle i|\hat{\tilde{V}}^{X,g}_{\text{sr-xc}}|0\rangle, \tag{42c}$$

$$\sigma^{X;o}_{\text{sr-xc},pq} = \tilde{g}^X_{\text{sr-xc},pq} + \langle 0|[\hat{E}^-_{pq}, \hat{V}^{X[1o]}_{\text{sr-xc}} + \hat{\tilde{V}}^{X,g}_{\text{sr-xc}}]|0\rangle + \frac{1}{2}\sum_t \left(g^X_{\text{sr-xc},tp} b^o_{qt} - g^X_{\text{sr-xc},tq} b^o_{pt} \right). \tag{42d}$$

The orbital gradient elements, $g^X_{\text{sr-xc},pq}$, were defined in Eqs. (24a) and (24b), the $\hat{V}^{X,g}_{\text{sr-xc}}$ operators were given in Eq. (25), and $\hat{\tilde{V}}^{X,g}_{\text{sr-xc}}$ is the one-index transformed form of the $\hat{V}^{X,g}_{\text{sr-xc}}$ operator in Eq. (25). The one-index transformed operators employ transformed integrals,[60,62] e.g., $\hat{\tilde{V}}^{X,g}_{\text{sr-xc}}$ is equivalent to Eq. (25), but with the $V^{X,g}_{\text{sr-xc},pq}$ integrals replaced by

$$\tilde{V}^{X,g}_{\text{sr-xc},pq} = \sum_t (V^{X,g}_{\text{sr-xc},tq} b^o_{pt} + V^{X,g}_{\text{sr-xc},pt} b^o_{qt}). \tag{43}$$

Similarly, the expression for the transformed gradient elements, $\tilde{g}^X_{\text{sr-xc},pq}$, is equivalent to Eqs. (24a) and (24b), but with $\hat{V}^{X,g}_{\text{sr-xc}}$ replaced with $\hat{\tilde{V}}^{X,g}_{\text{sr-xc}}$. The terms with $\hat{V}^{X,g}_{\text{sr-xc}}$ and $\hat{\tilde{V}}^{X,g}_{\text{sr-xc}}$ operators arise from the last terms of Eqs. (36) and (37), cf. the part within curly brackets. For these terms, the regular Hessian structure can be discerned.

At this point, it is noted that the $\boldsymbol{\sigma}^{S;c}$ and $\boldsymbol{\sigma}^{S;o}$ terms in Eqs. (42a)–(42d) enter exclusively for open-shell systems and are thus part of the extension for this work. Accordingly, with respect to the Hessian, this work concerns the implementation of the effective operator of triplet type, i.e., Eq. (41b), employed to define the $\boldsymbol{\sigma}^{S;c}$ and $\boldsymbol{\sigma}^{S;o}$ terms. In addition, the term that depends on $\mathbf{D}^{S(1\lambda)}$ in Eq. (41a) also enters exclusively for open-shell systems and has been implemented here.

In addition to the contributions from the exchange-correlation functional, the direct Hessian terms also contain contributions from the short-range Hartree term, $E_{\text{sr-H}}[\rho_C]$, in Eq. (14),

$$\frac{\partial^2 E_{\text{sr-H}}[\rho_C]}{\partial \lambda_j \partial \lambda_i} = \sum_{pq,rs} \frac{\partial D^C_{rs}(\boldsymbol{\lambda})}{\partial \lambda_j} g^{\text{sr}}_{pqrs} \frac{\partial D^C_{pq}(\boldsymbol{\lambda})}{\partial \lambda_i} + \sum_{pq,rs} D^C_{rs}(\boldsymbol{\lambda}) g^{\text{sr}}_{pqrs} \frac{\partial^2 D^C_{pq}(\boldsymbol{\lambda})}{\partial \lambda_j \partial \lambda_i}. \tag{44}$$

As for the short-range exchange-correlation contributions, the direct Hessian short-range Hartree contributions can be obtained by contraction of configurational and orbital trial vectors, respectively. The terms have the same structure as Eqs. (42a)–(42d), but without any spin-dependence, and can be defined employing the effective operator

$$\hat{V}^{C[1\lambda],c}_{\text{sr-H}} = \sum_{pq} j^{C(1\lambda)}_{\text{sr},pq}[\mathbf{D}^{C(1\lambda)}] \hat{E}_{pq}$$

with

$$j^{C(1\lambda)}_{\text{sr},pq}[\mathbf{D}^{C(1\lambda)}] = \sum_{rs} D^{C(1\lambda)}_{rs} g^{\text{sr}}_{rspq}. \tag{45}$$

We can combine the operators utilized for the short-range Hartree and short-range exchange-correlation terms in a common, effective operator

$$\hat{V}^{X[1\lambda]}_{\text{sr-Hxc}} = \begin{cases} \hat{V}^{C[1\lambda]}_{\text{sr-H}} + \hat{V}^{C[1\lambda]}_{\text{sr-xc}}, & X = C \\ \hat{V}^{S[1\lambda]}_{\text{sr-xc}}, & X = S \end{cases}, \tag{46}$$

which replaces $\hat{V}^{X[1\lambda]}_{\text{sr-xc}}$ in Eqs. (42a)–(42c).

III. IMPLEMENTATION

The configurational and orbital gradients for MC–srDFT were given in Eqs. (23)–(31) and depend on $\hat{V}^{C,g}_{\text{sr-xc}}$ and $\hat{V}^{S,g}_{\text{sr-xc}}$ operators [Eq. (25)]. The implementations of the equations for the gradient and (direct) Hessian in regular MCSCF within DALTON have been described previously.[58–60] A central part of this implementation builds on the generation of generalized Fock-type matrices

$$f_{pq} = \sum_r D^C_{pr}\, h_{qr} + \sum_{rst} P_{pr,st}\, g_{qr,st}, \tag{47}$$

where $\mathbf{P}$ is the reduced two-electron density matrix with elements $P_{pr,st} = \langle\Psi|\hat{e}_{pr,st}|\Psi\rangle$. In the following, we will in addition to the general orbitals (with indices p, q, r, s) need to denote both inactive and active orbitals, for which we use the indices i, j and u, v, x, y, respectively. In terms of the generalized Fock matrix, the orbital gradient becomes

$$g^o_{rs} = 2(f_{rs} - f_{sr}). \tag{48}$$

The direct Hessian is constructed in terms of modified gradient expressions, which in Ref. 60 are denoted by transition and one-index transformed gradient expressions. These can again be constructed from Eq. (47) by employing a one-index transformed density matrix or a transition density matrix instead of the reference density matrix.[60,63] The only non-zero long-range contributions to the Fock matrix elements are[60]

$$f^{\text{lr}}_{iq} = 2(f^{I,\text{lr}}_{iq} + f^{A,\text{lr}}_{iq}), \tag{49}$$

$$f^{\text{lr}}_{vq} = \sum_u D^C_{uv} f^{I,\text{lr}}_{uq} + Q^{\text{lr}}_{vq}, \tag{50}$$

where $Q^{\text{lr}}_{vq} = \sum_{uxy} P_{vu,xy}\, g^{\text{lr}}_{qu,xy}$. The Fock matrices $\mathbf{f}^{I,lr}$ and $\mathbf{f}^{A,lr}$ are called the *inactive* and *active* Fock matrix, respectively. They are defined as in Eqs. (3.11) and (3.12) of Ref. 60, but with long-range two-electron integrals. For an MC–srDFT description, the generalized Fock matrix elements must be altered to accommodate the potential contributions from the short-range DFT functional. Equations (49) and (50) become

$$f_{iq} = f^{\text{lr}}_{iq} + V^{C,g}_{\text{sr-Hxc},iq} + V^{S,g}_{\text{sr-xc},iq}, \tag{51}$$

$$f_{vq} = \sum_u D^c_{uv}\left(f^{I,\text{lr}}_{uq} + V^{C,g}_{\text{sr-Hxc},uq} + V^{S,g}_{\text{sr-xc},uq}\right) + Q^{\text{lr}}_{vq}, \tag{52}$$

where $V^{X,g}_{\text{sr-Hxc},pq}$ are the matrix elements of the operator in Eq. (30). Depending on the employed (spin-)density matrix, Eqs. (51) and (52) can be either a regular (spin-)Fock-matrix, a transition-(spin-)Fock matrix, or an one-index transformed (spin-)Fock matrix.

The auxiliary routines that construct Fock matrices were modified to accommodate short-range functionals in connection with previous work.[39] These auxiliary routines have been extended with the construction of the spin-density matrix required for the spin-dependent terms of $V^{X,g}_{\text{sr-Hxc},pq}$. The direct Hessian [Eqs. (42a)–(42d)] is comprised of terms that include $\mathbf{D}^{X(1\lambda)}$ (linear transformed terms) and "regular" Hessian terms. The former terms required slight modification for this work, whereas the latter terms are automatically included after the above-mentioned modifications of the gradient routines.

Finally, we have added code for the spin-dependent short-range exchange-correlation functionals from Ref. 48 and the spin-dependent gradient-corrected functional of Ref. 49 (see Sec. IV). This includes code for the numerical evaluation of the gradient and Hessian of the short-range functional kernel with respect to the spin density.

IV. COMPUTATIONAL DETAILS

The theory presented in Secs. II and III was implemented in a development version of the DALTON 2016 program.[46,47] All CAS–srDFT calculations were carried out with this new implementation. We have employed CAS–srDFT for two molecules, namely, O_2 and $[Fe(H_2O)_6]^{3+}$, exploiting point group symmetry: the D_{2h} symmetry sub-group of $D_{\infty h}$ for O_2 and the C_i symmetry group for $[Fe(H_2O)_6]^{3+}$.

For O_2, we focus on the $^3\Sigma^-_g$ and $^1\Delta_g$ lowest-energy states. The calculations on O_2 were carried out at the experimental ground-state equilibrium bond distance ($R_{\text{eq}} = 1.207$ Å) and at a stretched bond distance ($R_{\text{stretch}} = 2.0$ Å) with the cc-pVTZ basis set.[64,65] For O_2, we have employed three CAS(n, m) spaces with n electrons in m orbitals; CAS(8,6), CAS(12,8), and CAS(12,16). These spaces correspond to including the orbitals $3\sigma_g^2 1\pi_u^4 1\pi_g^{*2} 3\sigma_u^{*0}$, $2\sigma_g^2 2\sigma_u^{*2} 3\sigma_g^2 1\pi_u^4 1\pi_g^{*2} 3\sigma^{*0}$, and $2\sigma_g^2 2\sigma_u^{*2} 3\sigma_g^2 1\pi_u^4 1\pi_g^{*2} 3\sigma_u^{*0} 4\sigma_g^0 4\sigma_u^{*0} 5\sigma_g^0 2\pi_u^0 2\pi_g^{*0} 5\sigma_u^{*0}$, respectively, where the occupations are for the reference configuration.

For $[Fe(H_2O)_6]^{3+}$, we focus on the ^6A_g and $^4T_{1g}$ lowest-energy states (using a nomenclature that assumes octahedral symmetry), and we used the structure from Ref. 66, which was optimized for the ^6A_g ground state within a cluster including the second solvation shell. We have removed this second solvation shell, and our calculations only concern the $[Fe(H_2O)_6]^{3+}$ core. Note that the CASPT2 value from Ref. 66 to which we compare was obtained exactly this way (cf. Table 7 of Ref. 66). It was shown that such a calculation includes a large part of the solvation effect. For this complex, we employed the def2-TZVPP basis set[67] for iron and oxygen and the def2-SV(P) basis set for hydrogen. Following Ref. 66, we employed an active space that includes the five d orbitals, two additional σ Fe–O bonding orbitals, and a second shell of d orbitals yielding a CAS(9,12) active space. The orbitals comprising the active space are shown in Fig. 1.

As short-range exchange-correlation functionals we employed the spin-dependent srLDA functional from Ref. 48 and the spin-dependent gradient-corrected functional from Ref. 49, denoted by srPBE. The name of the latter reflects that it originates from the Perdew-Burke-Ernzerhof (PBE) functional[68] and has been extended to the short-range interaction in Ref. 69 and further modified by Goll *et al.*[44,49]

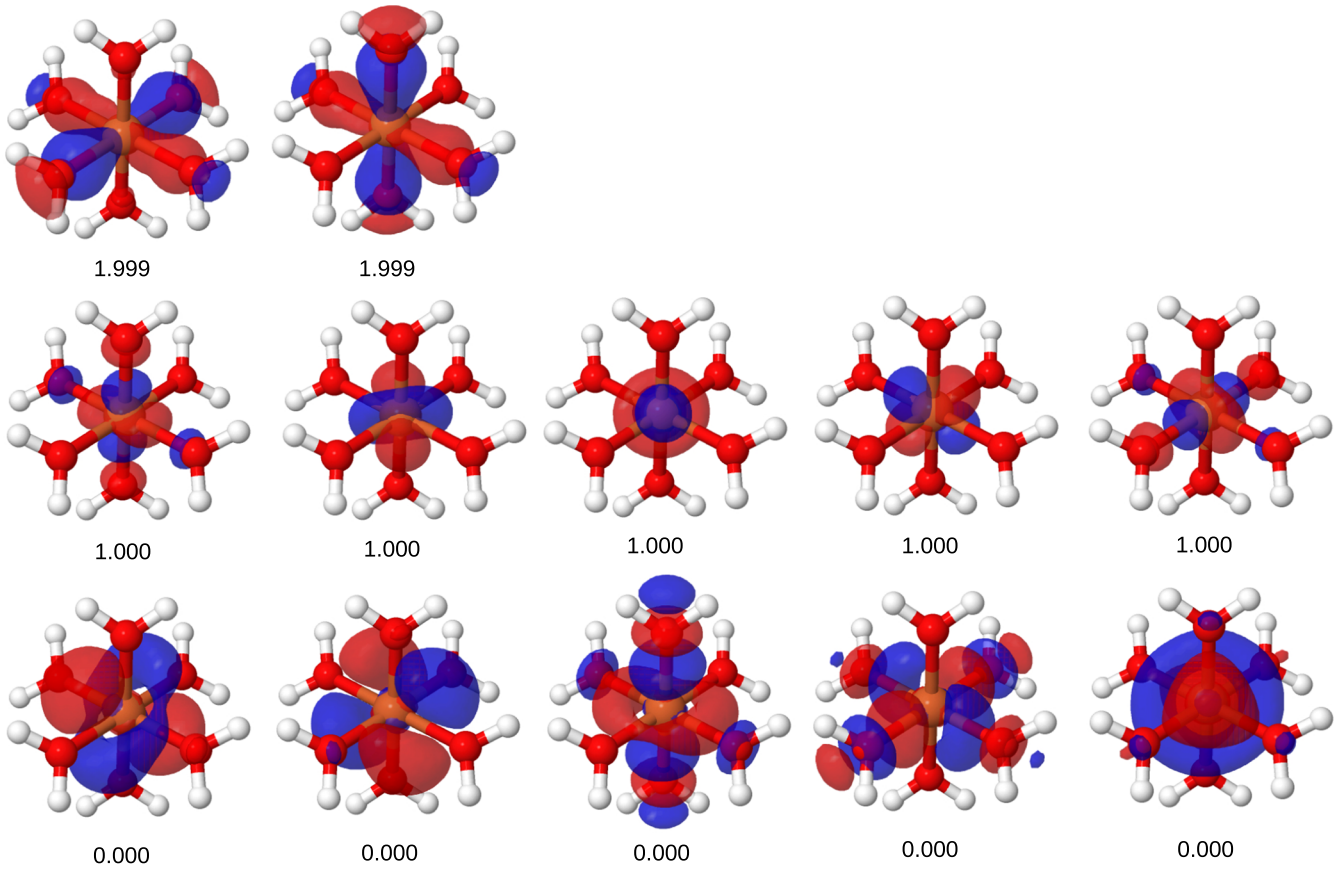

FIG. 1. Active orbitals for the sextet state of $[Fe(H_2O)_6]^{3+}$. Occupation numbers are shown below (a similar active space was employed for the quartet state).

V. DEMONSTRATION OF THE OPEN-SHELL CAS–srDFT METHOD

We here discuss the results from CAS-srLDA and CAS-srPBE calculations on two prototype open-shell systems.

A. The dioxygen molecule

For dioxygen, we start with a discussion of the optimal value of the range-separation parameter μ. This matter has been discussed for CAS–srDFT[27,35] and also for other range-separated methods. Note that with an exact srDFT functional and a full CI long-range wave function (in a complete basis) the exact result would be obtained independent of the μ value. The μ value determines the mixture between the wave function and srDFT parts and is thus a parameter that shifts correlation effects between the two parts. When we refer to an "optimal μ value" in the following, we mean the value giving accurate results with the least possible computational effort. As the computational time for the srDFT part is basically independent of μ while the time for accurate long-range wave function contributions grows rapidly with μ, the optimal μ value is the smallest μ value for which the non-local correlations not treated properly by semi-local srDFT functionals are *not* in the realm of the srDFT part. For multireference models, one recipe has been to compare HF–srDFT and CAS–srDFT. From this, it was found that a value of μ of 0.4 bohr^{-1} allocates the main part of the dynamical correlation to the srDFT functional, while the long-range wave function still incorporates a substantial part of static correlation in multiconfigurational systems.[27] The $\mu = 0.4$ value was recently confirmed for excitation energies obtained from linear-response theory.[35] Here, we investigate the optimal value of μ for the O_2 molecule in its triplet ground state $^3\Sigma_g^-$. It is expected that this state is dominated by the electronic configuration

$$\begin{aligned} |\Psi^{O_2}_{R_{eq}}(^3\Sigma_g^-)\rangle &= C_0\ {}^3|1\sigma_g^2 1\sigma_u^{*2} 2\sigma_g^2 2\sigma_u^{*2} 3\sigma_g^2 1\pi_{u,x}^2 1\pi_{u,y}^2 1\pi_{g,x}^{*1} 1\pi_{g,y}^{*1} 3\sigma_u^{*0}\rangle + \cdots \\ &\equiv C_0\ {}^3|\pi_{u,x}^2 \pi_{u,y}^2 \pi_{g,x}^{*1} \pi_{g,y}^{*1}\rangle + \cdots \end{aligned} \tag{53}$$

at the equilibrium bond distance. The $^3|\ldots\rangle$ notation includes the spin multiplicity of the configuration state function, and we have in the last equality left out the σ electrons for brevity. This dominant configuration can be described by a single-determinant high-spin wave function ($M_S = 1$), which means that HF–srDFT and CAS–srDFT should provide

similar energies. In order to have a system that displays multiconfigurational character and is directly comparable to the above system, we also consider a bond distance corresponding to a stretched bond (R_{stretch} = 2.0 Å). In this case, we expect a splitting of HF–srDFT and CAS–srDFT energies, and the parameter μ should be chosen to reflect this.

The HF–srDFT and CAS–srDFT energies as functions of the value of μ are shown in Fig. 2, both for R_{eq} and R_{stretch}. For the equilibrium bond distance, the HF–srDFT and MC–srDFT energies are identical (within 10^{-3} hartree) until we reach a value of $\mu \approx 0.4$. This value is obtained for both of the employed srDFT functionals, and the result is thus very similar to what was found in Ref. 27. We have tested three different active spaces of increasing size. The smallest CAS(8,6) contains only σ and π orbitals from the oxygen 2p orbitals, whereas the CAS(12,8) active space additionally includes the σ orbitals with origin in the 2s orbitals and thus constitutes the full valence space. Finally, the CAS(12,16) is an extended active space also including the oxygen 3s and 3p orbitals. For R_{eq}, the splitting occurs at around the same value of μ (cf. Fig. 2, left), independent of the employed active space.

For an elongated bond, the system becomes multiconfigurational as can be seen from the fact that the HF–srDFT and MC–srDFT energies are clearly different, already at small values of μ (cf. Fig. 2, right). Thus, a value of $\mu = 0.4$ also allows the long-range CAS wave function to include static correlation for the present multiconfigurational open-shell system.

We further comment on a few more observations from Fig. 2. First, we note that for the srPBE functional, the μ dependence is benign for small μ values in the case of O_2. For this small molecule, the long-range CASSCF wave function can cover the correlation up to about $\mu = 0.6$–0.7, as seen from the almost horizontal energy curves in this range of values of μ. As expected, this is in contrast with the CAS–srLDA energy, which is too high for $\mu = 0$ and therefore decreases with increasing μ since the long-range wave function with increasing μ can describe an increasing part of the correlation. That this happens *before* the HF-srLDA and MC-srLDA curves separate must mean that this effect is mostly in the exchange part. Furthermore, that the CAS-srLDA energies decrease with increasing μ while CAS-srPBE energies increase indicates that pure CASSCF is better than any of the CAS-srLDA models, while CAS-srPBE contains some correlation effects not described by the pure CASSCF.

It is also interesting to study the first singlet excited state $^1\Delta_g$. In D_{2h} symmetry, this electronic state can be characterized either by a dominant configuration given by a linear combination of the Slater determinants in which the π^*_g orbitals are doubly occupied, i.e.,

$$|\Psi^{O_2}_{R_{\text{eq}}}(^1\Delta_g^{x^2-y^2})\rangle = C_0 \frac{1}{\sqrt{2}}(^1|\pi^2_{u,x}\pi^2_{u,y}\pi^{*2}_{g,x}\pi^{*0}_{g,y}\rangle - {}^1|\pi^2_{u,x}\pi^2_{u,y}\pi^{*0}_{g,x}\pi^{*2}_{g,y}\rangle) + \cdots, \quad (54)$$

or by a dominant configuration in which the $\pi^*_{g,x}$ and $\pi^*_{g,y}$ orbitals are spin-singlet coupled,

$$|\Psi^{O_2}_{R_{\text{eq}}}(^1\Delta_g^{xy})\rangle = C_0\, {}^1|\pi^2_{u,x}\pi^2_{u,y}\pi^{*1}_{g,x}\pi^{*1}_{g,y}\rangle + \cdots. \quad (55)$$

The states in Eqs. (54) and (55) are exactly degenerate in $D_{\infty h}$ but belong to different irreducible representations in the D_{2h} subgroup: A_{1g} and B_{1g}. None of the dominant configurations in these two wave functions can be represented by a spin-restricted single determinant and thus cannot be obtained in standard spin-restricted Kohn-Sham DFT. By contrast, the long-range wave function of MC–srDFT can describe the $^1\Delta_g$ state in both the D_{2h} irreducible representations of Eqs. (54) and (55). The method correctly predicts that the $^1\Delta_g^{x^2-y^2}$ and $^1\Delta_g^{xy}$ states are degenerate. As expected, the dominant configurations in the long-range CAS wave functions

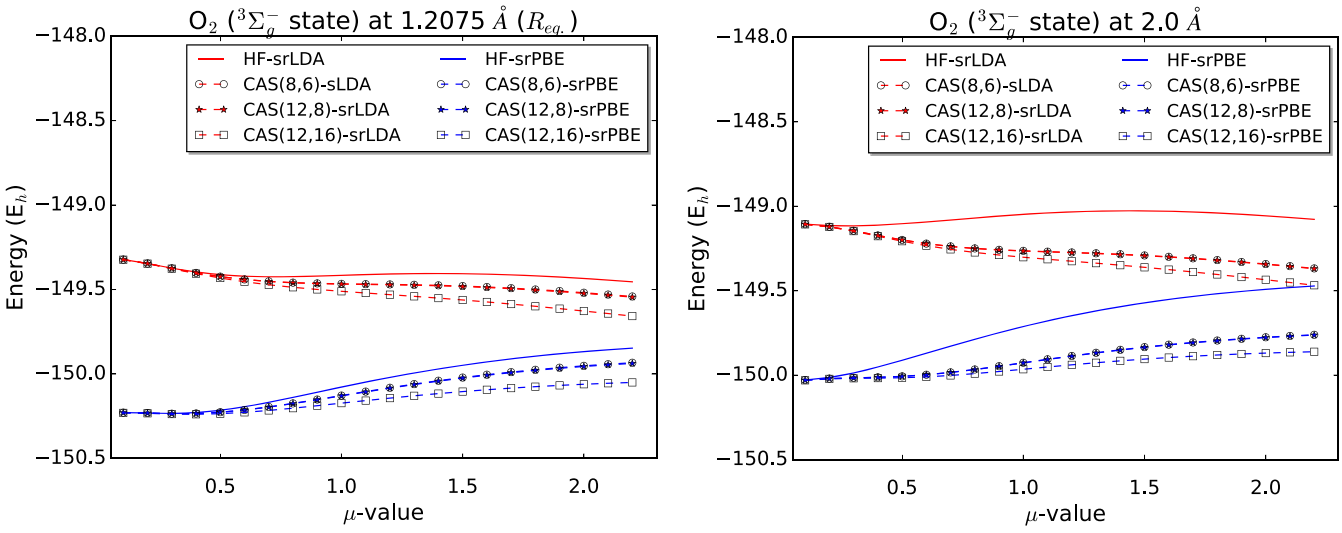

FIG. 2. Dependence on the range-separation parameter μ of the HF–srDFT and CAS–srDFT energies for O_2 in its $^3\Sigma_g^-$ ground state for the equilibrium bond distance R_{eq} (left) and a stretched bond distance R_{stretch} (right). All calculations are with the cc-pVTZ basis set.

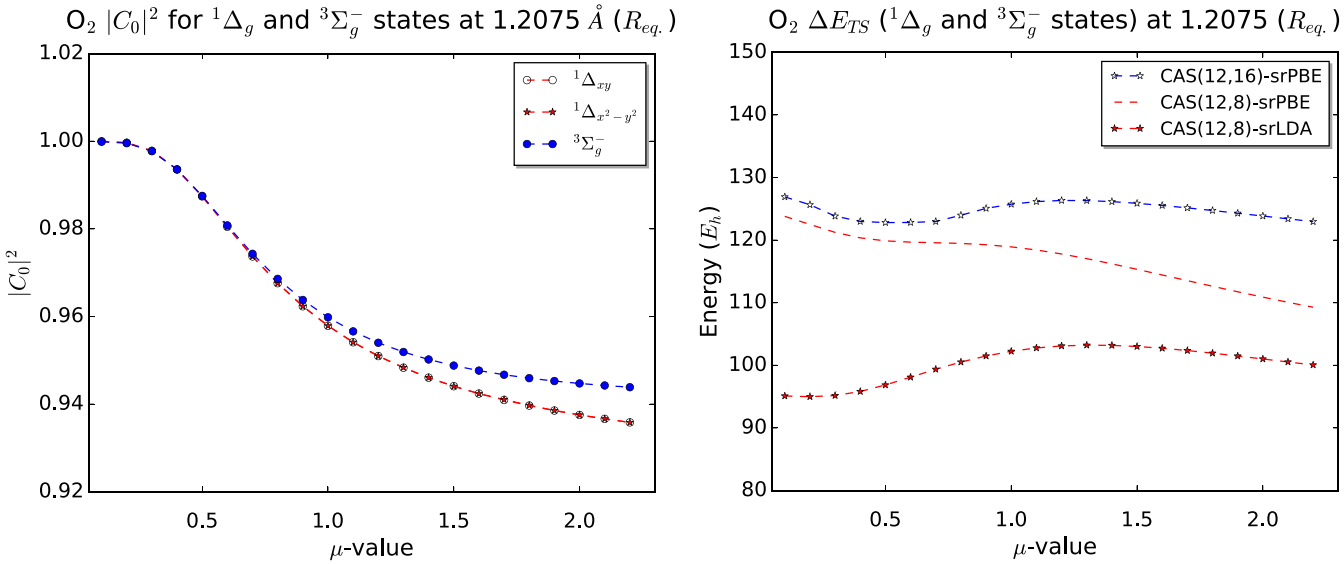

FIG. 3. Left: Weight $|C_0|^2$ of the dominant configuration in the long-range CAS(12,8) wave function for O_2 in the $^3\Sigma_g^-$ and $^1\Delta_g$ states at the equilibrium bond distance. The srLDA and srPBE functionals give rise to practically identical long-range wave functions. Right: Singlet-triplet splitting $\Delta E(^3\Sigma_g^- \rightarrow {}^1\Delta_g)$.

are indeed given as the ones in Eqs. (54) and (55). In Fig. 3 (left), we have reported the value of $|C_0|^2$ for the dominant configuration in the long-range wave function as a function of μ. The $^3\Sigma_g^-$, $^1\Delta_g^{x^2-y^2}$, and $^1\Delta_g^{xy}$ states behave remarkably similarly, in agreement with textbook qualitative MO theory, and the same trend is observed: a steadily decreasing value of $|C_0|^2$ as we increase μ, reflecting that the long-range wave function must become more multiconfigurational to treat the electron correlation effects moved from the short-range to the long-range parts with increasing μ. The value of $|C_0|^2$ for the $^1\Delta$ states decreases fastest with increasing mixture of long-range wave function. Yet, at small values of μ (including $\mu = 0.4$), we can expect that the singlet-triplet splitting will be determined to a large degree by the leading configuration for each state and by how the short-range functionals translate the density obtained from the leading configuration into an energy. Any possible error is therefore likely due to the srDFT functional (as will be further discussed below). As a remark, we note that the $^1\Delta_g^{x^2-y^2}$ and $^1\Delta_g^{xy}$ energies as well as their respective $|C_0|^2$ values are indistinguishable at all values of μ. This shows that the method treats singlet-coupled open shells qualitatively correct.

With these results in mind, we can also investigate the singlet-triplet splitting $\Delta E(^3\Sigma_g^- \rightarrow {}^1\Delta_g)$. We note, however, that we cannot expect accurate results due to deficiencies in the employed short-range functionals. In particular, the short-range exchange does not correctly cancel the short-range Hartree self-repulsion for the singly occupied π orbitals in singlet states. In the case of O_2, this can be seen from the above analysis of the dominant configurations: the long-range wave function correctly describes a spin-singlet coupled state in which the total spin density is zero but which locally may become both positive and negative [perhaps as shown most obviously in the $^1\Delta_g^{xy}$ state, Eq. (55)]. Yet, our present srDFT functionals are not equipped with the ability to employ local spin densities, as required for a spin-singlet coupled state. We therefore expect too high energy of the $^1\Delta_g$ state and thus of the splitting (as indeed observed, see below). Our investigation of the singlet-triplet splitting therefore focus mainly on the μ-dependence, whereas the values itself are only briefly discussed. The dependence on μ of the singlet-triplet splitting is shown in Fig. 3 (right) for different active spaces and short-range functionals and given in Table I for selected values of μ with CAS(12,8)–srDFT and CAS(12,16)–srDFT. In general, the dependence on μ is moderate (at most a few kJ/mol) in the interval $\mu = 0.1$–0.5. The difference is larger (≈25–30 kJ/mol) between the results obtained with the srLDA and srPBE functionals; thus there is a significant effect

TABLE I. Vertical singlet-triplet splitting, $\Delta E(^3\Sigma_g^- \rightarrow {}^1\Delta_g)$ (in kJ/mol), for O_2 at the equilibrium bond distance, calculated with CAS–srDFT employing different srDFT functionals and values of μ (in bohr^{-1}).

Method	$\mu = 0.0$	$\mu = 0.1$	$\mu = 0.3$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 1.0$	$\mu \rightarrow \infty$
CAS(12,8)-srLDA	96.96	95.11	95.18	95.82	96.86	102.24	92.35
CAS(12,16)-srLDA	96.96	96.97	97.66	100.13	103.53	107.78	110.44
CAS(12,8)-srPBE	127.62	123.79	121.24	120.36	119.88	118.92	92.75
CAS(12,16)-srPBE	127.62	126.91	123.83	122.99	122.80	125.74	110.44
Expt.[70]				94.6			

TABLE II. Singlet-triplet splitting, $\Delta E(^3\Sigma_g^- \rightarrow {}^1\Delta_g)$ (in kJ/mol), for O_2 at the equilibrium bond distance. The CAS-srLDA and CAS-srPBE results are for $\mu = 0.4$ bohr^{-1}. Note that Refs. 52 and 53 report the adiabatic energy difference, but with a difference of 0.008 Å between R_e for the two states, this is expected to be of minor importance.

Method	$\Delta E(^3\Sigma_g^- \rightarrow {}^1\Delta_g)$
Our work (all with cc-pVTZ)	
CAS(12,8)	92.35
CAS(12,8)-srLDA	95.82
CAS(12,8)-srPBE	120.36
CAS(12,16)	110.44
CAS(12,16)-srLDA	100.13
CAS(12,16)-srPBE	122.99
Literature	
CCSD(T)/cc-pVTZ[52]	124.66
CCSDT/cc-pVTZ[52]	114.34
CCSDTQ/cc-pVTZ[52]	99.48
MRCI[51]	111.61
MRCI/cc-pVTZ[52]	98.40
QMC[53]	96.58
Expt.[70]	94.6

of employing a gradient-corrected functional. Although the srLDA results are in fact closer to the experimental value than the srPBE results, both functionals have the above-mentioned flaws leading to destabilization of the singlet energy. Therefore, the good performance of srLDA must be considered fortuitous. We compare the splitting obtained with srLDA and srPBE to other results from the literature in Table II. Our best theoretical CAS(12,16)–srPBE result yields a splitting of 123 kJ/mol, which is in between the CCSD(T) and CCSDT results, but (as expected) too large compared to the experimental value of 94.6 kJ/mol.[70] The fact that even CCSDT and some of the MRCI results[51] overestimate the experiment value by 20 and 16 kJ/mol, respectively, emphasizes that the calculation of this singlet-triplet splitting energy is indeed delicate. Only the highly accurate CCSDTQ and QMC methods are able to achieve a result that is within 5 kJ/mol of the experimental value. A peculiar observation is that the small active space CAS(12,8) is very close to the experimental result, and CAS(12,8)-srLDA is almost identical to experiment. Both results are due to fortuitous error cancellation, as extending the active space leads to deterioration. Yet, CAS(12,16) and CAS(12,16)-srLDA are in fact still closer to experiment than both CCSDT and MRCI, showing that assessment of the functionals can require some consideration of the active space employed. Finally, we note that approaches based on perturbation theory (e.g., CASPT2 or NEVPT2) are likely also to provide a reasonable spin-state splitting. Indeed, we performed an NEVPT2(12,8) calculation which gives a singlet-triplet splitting of 92.35 kJ/mol. However, an in-depth discussion of CASPT2/NEVPT2 results is out of the scope for our current paper, where we focus on MC–srDFT. To improve spin-state energetics, we are planning to make a number of improvements on the srDFT functionals, e.g., employing local spin densities. One option could be to employ the on-top pair density[23,71] for this purpose.

B. The $^6A_g \rightarrow {}^4T_{2g}$ spin-state splitting in $[Fe(H_2O)_6]^{3+}$

We now turn to a system where methods such as CCSDTQ are probably beyond reach. The prototypical $[Fe(H_2O)_6]^{3+}$ complex has been described in numerous theoretical studies with a variety of methods. We will here focus on the $^6A_g \rightarrow {}^4T_{2g}$ transition, where in our labeling we assume octahedral symmetry. The higher spin state (^6A_g) can qualitatively be described as an open-shell d-complex with five singly occupied d-orbitals, whereas the lower spin state ($^4T_{2g}$) can be described as a doubly occupied orbital in the lowest-lying t_{2g} orbital level and a hole in the e_g level. The lower spin-state thus does not involve a coupling between spatially separated orbitals, suggesting that we will not observe errors due to the lack of local spin dependence in the employed srDFT functionals as described in Sec. V A.

The experimental band for the $^6A_g \rightarrow {}^4T_{2g}$ transition has a maximum at 12 600 cm^{-1} (1.56 eV) in acidic, aqueous solution.[72] However, previous results obtained with CASPT2 and CCSD(T) have showed remarkably large deviations (up to 1.5 eV) from this value.[54,73] Large deviations were also found with the spectroscopically oriented configuration interaction (SORCI) method.[74] On the other hand, both DFT[75–77] and the semi-empirical INDO/S method seem to be reasonably accurate.[75,78] A recent study by Radoń *et al.*[66] pointed out some of the possible failures of accurate *ab initio* methods. First, it was argued that the second solvation shell was important, mainly for the molecular structure of the $[Fe(H_2O)_6]^{3+}$ core. If this effect is taken into account, even calculations on $[Fe(H_2O)_6]^{3+}$ are in significant better agreement with experiment. Further, Radoń *et al.*[66] argued that in addition to the *d* orbitals, the employed CAS space should at least contain two σ Fe–O bonding orbitals.[73] This leads to a CAS space of 9 electrons in 12 orbitals, whereas some earlier studies employed the smaller CAS(5,5) and CAS(5,10) active spaces.[73] The effect was estimated to be around 0.24 eV on the transition energy. However, Radoń *et al.*[66] also stressed that their study was not directly comparable to some of the previous studies as they (unlike Ref. 73) employed an IPEA-shifted zeroth-order Hamiltonian and also larger basis sets of which the combined effect can be estimated to be around 0.3 eV on the transition energy.

Here, we start by investigating the dependence on μ of the spin-state splitting $\Delta E(^6A_g \rightarrow {}^4T_{1g})$. Results for selected values of μ are shown in Table III. Compared to the spin-state

TABLE III. Vertical quartet-sextet splitting (in eV) of $[Fe(H_2O)_6]^{3+}$, calculated with CAS–srPBE employing different values of μ (in bohr^{-1}).

Method	$\mu = 0.01$	$\mu = 0.1$	$\mu = 0.33$	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.75$	$\mu = 1.0$	$\mu = \infty$
CAS(9,12)-srPBE	1.15	1.15	1.35	1.44	1.61	1.74	2.10	2.24

splitting in O_2 (see Sec. V A), the $\Delta E(^6A_g \rightarrow {}^4T_{1g})$ spin-state splitting in the range $\mu = 0.4$–1.0 bohr^{-1} varies somewhat more with 1.15–2.10 eV for CAS(9,12)-srPBE. For values in the interval $\mu = 0.33$ to $\mu = 0.50$ bohr^{-1} expected *a priori* to be optimal, we do obtain good agreement with the experimental value of 1.56 eV. The "end-points" in Table III corresponds to a pure CAS(9,12) ($\mu = \infty$) and pure PBE ($\mu = 0.01$); these, respectively, overestimate and underestimate the experimental value significantly. It is gratifying the range of μ-values expected to provide the best results, indeed all are the best ones obtained (and, in fact, all closer to experiment than CASPT2 with srPBE). Yet, before more general conclusions concerning the performance can be made, a larger span of systems must be explored. In Table IV, we compare the CAS(9,12)-srPBE result at $\mu = 0.4$ bohr^{-1} with the CAS(9,12)-srLDA one, as well as results obtained with a number of other theoretical methods. As mentioned above, the literature results show a rather large scatter, and all the values for the quartet-sextet splitting are too large. Our result with CAS(9,12)-srLDA shows the opposite pattern; it underestimates the experimental value significantly. The corresponding values for $\mu = 0.33$ and $\mu = 0.5$ are 1.04 eV and 1.29 eV and are thus still significantly underestimated. However, the gradient-corrected srDFT functional significantly improves the spin-state splitting energy for the value $\mu = 0.4$. This shows that a GGA srDFT functional is important, as it also was for O_2. We found in Sec. V A that $\mu = 0.4$ was a good compromise between flexibility of the long-range wave function to include static correlation and ability of the srDFT functional to recover dynamical correlation. This value also provides a very accurate estimate for the quartet-sextet splitting $\Delta E(^6A_g \rightarrow {}^4T_{1g})$ with CAS(12,9)-srPBE. The result is 1.44 eV, only 0.12 eV below the experimental value. In comparison, a calculation with CAS(9,12)PT2 on the same $[Fe(H_2O)_6]^{3+}$ molecular structure gives a spin-state splitting of 1.96 eV, which is 0.44 eV too high. We note that $\mu = 0.33$ and $\mu = 0.5$ also give splittings better than the CAS(9,12)PT2 values.

TABLE IV. Vertical quartet-sextet splitting (in eV) for $[Fe(H_2O)_6]^{3+}$. The CASPT2, CAS(9,12)-srLDA, and CAS(9,12)-srPBE calculations are for the $[Fe(H_2O)_6]^{3+}$ core with a molecular structure that included the second solvation shell in the optimization; see Ref. 66. The CAS(9,12)-srLDA and CAS(9,12)-srPBE calculations are for a range-separation parameter of $\mu = 0.4$ bohr^{-1}.

Method	$\Delta E(^6A_g \rightarrow {}^4T_{1g})$
$[Fe(H_2O)_6]^{3+}$ optimized in vacuum	
Literature	
CCSD(T)[54]	2.47
SORCI[74]	2.07
CASPT2[54]	2.64
$[Fe(H_2O)_6]^{3+}$ optimized with a second solvation sphere	
Literature	
CASPT2[66]	1.96
Our work	
CAS(9,12)-srLDA	1.14
CAS(9,12)-srPBE	1.44
Expt.[72]	1.56

VI. CONCLUSIONS

In this paper, we have presented the theory and implementation of a generalization of the MC–srDFT method to employ spin-dependent short-range density functionals. We have applied the method to two well-known cases, namely, the O_2 molecule which has a triplet ground state and the transition metal complex $[Fe(H_2O)_6]^{3+}$.

For the O_2 molecule, a wealth of literature data exists for the spin-state splitting between the ${}^3\Sigma_g^-$ ground state and the lowest-lying singlet state ${}^1\Delta_g$, and QMC or CCSDTQ methods provide a value in very good agreement with experiment. The spin-state splitting has here been investigated with a range of values for the range-separation parameter μ, but the splitting is not strongly dependent on μ. Compared to the experimental value, CAS–srPBE lies between CCSD(T) and CCSDT in accuracy. This is already impressive, but there is also room for improvement, as the ${}^1\Delta_g$ state is estimated too high in energy by both CCSD(T), CCSDT, and CAS-srDFT. The former two methods can be improved by higher-order cluster expansions, but this comes at a high computational cost. An improvement (that will be computationally cheap) for CAS-srDFT is to allow dependence of local spin densities in the employed srDFT functionals, which is expected to lower the energy of the spin-coupled ${}^1\Delta_g$ state.

Next, we investigated the spin-state splitting $\Delta E(^6A_g \rightarrow {}^4T_{1g})$ in $[Fe(H_2O)_6]^{3+}$. In this case, CCSD(T) and CASPT2 are both significantly off. A method such as CCSDTQ is too computationally demanding, while CAS–srPBE shows very good agreement with experiment.

It should also be emphasized that most other high-level methods require large basis set expansions with high-order angular momenta to describe the electron-electron Coulomb cusp accurately. The MC–srDFT method avoids this by replacing an explicit description of the Coulomb cusp with an effective density functional. Hence, the basis set requirements can be expected to be similar to regular DFT, which is known for its fast convergence with respect to the basis set expansion.

The first studies shown here are promising, but it should be emphasized that further developments are required. An obvious extension is to employ srDFT functionals that include the kinetic energy density, which has been developed for range-separated methods.[79] Also, a good measure of local spin density, for example, in terms of the on-top pair density,[71] needs to be implemented to satisfactorily describe, for example, open-shell singlets which exhibit zero mean spin density. Extension of the open-shell algorithm to molecular properties via response theory will also be an important extension of the method. Development in these directions is ongoing.

In the future, it would also be interesting to extend our state-averaged version of MC-srDFT,[32] in particular, for the unbiased location of conical intersections.

ACKNOWLEDGMENTS

E.D.H. thanks the Carlsberg Foundation and the European Commission for post-doc stipends. J.T. and H.J.A.J. acknowledge support from the French government via the French-Danish collaboration project "WADEMECOM.dk."

APPENDIX: GRADIENT CONTRIBUTIONS FROM SHORT-RANGE GGA FUNCTIONALS

For an srGGA functional, the short-range exchange-correlation energy density $e_{\text{sr-xc}}$ also depends on the three gradient variables $\xi_{CC}(\mathbf{r}) = \nabla\,\rho_C(\mathbf{r})\cdot\nabla\,\rho_C(\mathbf{r})$, $\xi_{SS}(\mathbf{r}) = \nabla\,\rho_S(\mathbf{r})\cdot\nabla\,\rho_S(\mathbf{r})$, and $\xi_{CS}(\mathbf{r}) = \nabla\,\rho_C(\mathbf{r})\cdot\nabla\,\rho_S(\mathbf{r})$ in addition to the electron charge density $\rho_C(\mathbf{r})$ and the electron spin density $\rho_S(\mathbf{r})$,

$$E_{\text{sr-xc}}[\rho_C,\rho_S,\xi_{CC},\xi_{SS},\xi_{CS}] = \int e_{\text{sr-xc}}(\rho_C(\mathbf{r}),\rho_S(\mathbf{r}),\xi_{CC}(\mathbf{r}),\xi_{SS}(\mathbf{r}),\xi_{CS}(\mathbf{r}))\,\mathrm{d}\mathbf{r}, \tag{A1}$$

where we for brevity define the integrand as $e_{\text{sr-xc}}$. The derivatives required for the electronic gradient then become

$$\frac{\partial e_{\text{sr-xc}}}{\partial\lambda_i} = \frac{\partial e_{\text{sr-xc}}}{\partial\rho_C}\frac{\partial\rho_C(\mathbf{r})}{\partial\lambda_i} + \frac{\partial e_{\text{sr-xc}}}{\partial\rho_S}\frac{\partial\rho_S(\mathbf{r})}{\partial\lambda_i} + \frac{\partial e_{\text{sr-xc}}}{\partial\xi_{CC}}\frac{\partial\xi_{CC}(\mathbf{r})}{\partial\lambda_i} + \frac{\partial e_{\text{sr-xc}}}{\partial\xi_{SS}}\frac{\partial\xi_{SS}(\mathbf{r})}{\partial\lambda_i} + \frac{\partial e_{\text{sr-xc}}}{\partial\xi_{CS}}\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_i} \tag{A2}$$

and the derivatives required for the electronic Hessian become

$$\begin{aligned}\frac{\partial^2 e_{\text{sr-xc}}}{\partial\lambda_i\partial\lambda_j} &= \sum_{X=C,S}\left(\frac{\partial e_{\text{sr-xc}}}{\partial\rho_X}\frac{\partial^2\rho_X(\mathbf{r})}{\partial\lambda_i\partial\lambda_j} + \frac{\partial e_{\text{sr-xc}}}{\partial\xi_{XX}}\frac{\partial^2\xi_{XX}(\mathbf{r})}{\partial\lambda_i\partial\lambda_j}\right) + \frac{\partial e_{\text{sr-xc}}}{\partial\xi_{CS}}\frac{\partial^2\xi_{CS}(\mathbf{r})}{\partial\lambda_i\partial\lambda_j}\\ &+ \sum_{X,Y=C,S}\left(\frac{\partial^2 e_{\text{sr-xc}}}{\partial\rho_Y\partial\rho_X}\frac{\partial\rho_X(\mathbf{r})}{\partial\lambda_i}\frac{\partial\rho_Y(\mathbf{r})}{\partial\lambda_j} + \frac{\partial^2 e_{\text{sr-xc}}}{\partial\rho_Y\partial\xi_{XX}}\frac{\partial\xi_{XX}(\mathbf{r})}{\partial\lambda_i}\frac{\partial\rho_Y(\mathbf{r})}{\partial\lambda_j}\right.\\ &\left.+ \frac{\partial^2 e_{\text{sr-xc}}}{\partial\xi_{YY}\partial\rho_X}\frac{\partial\rho_X(\mathbf{r})}{\partial\lambda_i}\frac{\partial\xi_{YY}(\mathbf{r})}{\partial\lambda_j} + \frac{\partial^2 e_{\text{sr-xc}}}{\partial\xi_{YY}\partial\xi_{XX}}\frac{\partial\xi_{XX}(\mathbf{r})}{\partial\lambda_i}\frac{\partial\xi_{YY}(\mathbf{r})}{\partial\lambda_j}\right)\\ &+ \sum_{X=C,S}\left(\frac{\partial^2 e_{\text{sr-xc}}}{\partial\xi_{CS}\partial\rho_X}\frac{\partial\rho_X(\mathbf{r})}{\partial\lambda_i}\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_j} + \frac{\partial^2 e_{\text{sr-xc}}}{\partial\xi_{CS}\partial\xi_{XX}}\frac{\partial\xi_{XX}(\mathbf{r})}{\partial\lambda_i}\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_j}\right) + \frac{\partial^2 e_{\text{sr-xc}}}{\partial\xi_{CS}^2}\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_i}\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_j}.\end{aligned} \tag{A3}$$

The first-order derivatives of the three gradient terms are

$$\frac{\partial\xi_{CC}(\mathbf{r})}{\partial\lambda_i} = 2\sum_{pq}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\rho_C(\mathbf{r}))\frac{\partial D^C_{pq}}{\partial\lambda_i}, \tag{A4}$$

$$\frac{\partial\xi_{SS}(\mathbf{r})}{\partial\lambda_i} = 2\sum_{pq}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\rho_S(\mathbf{r}))\frac{\partial D^S_{pq}}{\partial\lambda_i}, \tag{A5}$$

$$\frac{\partial\xi_{CS}(\mathbf{r})}{\partial\lambda_i} = \sum_{pq}\nabla\Omega_{pq}(\mathbf{r})\cdot\left(\nabla\rho_C(\mathbf{r})\frac{\partial D^S_{pq}}{\partial\lambda_i} + \nabla\rho_S(\mathbf{r})\frac{\partial D^C_{pq}}{\partial\lambda_i}\right), \tag{A6}$$

and the second-order derivatives are

$$\begin{aligned}\frac{\partial^2\xi_{CC}(\mathbf{r})}{\partial\lambda_i\partial\lambda_j} &= 2\sum_{pq,rs}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\Omega_{rs}(\mathbf{r}))\frac{\partial D^C_{rs}}{\partial\lambda_j}\frac{\partial D^C_{pq}}{\partial\lambda_i}\\ &+ 2\sum_{pq}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\rho_C(\mathbf{r}))\frac{\partial^2 D^C_{pq}}{\partial\lambda_i\partial\lambda_j},\end{aligned} \tag{A7}$$

$$\begin{aligned}\frac{\partial^2\xi_{SS}(\mathbf{r})}{\partial\lambda_i\partial\lambda_j} &= 2\sum_{pq,rs}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\Omega_{rs}(\mathbf{r}))\frac{\partial D^S_{rs}}{\partial\lambda_j}\frac{\partial D^S_{pq}}{\partial\lambda_i}\\ &+ 2\sum_{pq}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\rho_S(\mathbf{r}))\frac{\partial^2 D^S_{pq}}{\partial\lambda_i\partial\lambda_j},\end{aligned} \tag{A8}$$

$$\begin{aligned}\frac{\partial^2\xi_{CS}(\mathbf{r})}{\partial\lambda_i\partial\lambda_j} &= \sum_{pq,rs}(\nabla\Omega_{pq}(\mathbf{r})\cdot\nabla\Omega_{rs}(\mathbf{r}))\\ &\times\left(\frac{\partial D^C_{rs}}{\partial\lambda_j}\frac{\partial D^S_{pq}}{\partial\lambda_i} + \frac{\partial D^S_{rs}}{\partial\lambda_j}\frac{\partial D^C_{pq}}{\partial\lambda_i}\right) + \sum_{pq}\nabla\Omega_{pq}(\mathbf{r})\\ &\cdot\left(\nabla\rho_C(\mathbf{r})\frac{\partial^2 D^S_{pq}}{\partial\lambda_i\partial\lambda_j} + \nabla\rho_S(\mathbf{r})\frac{\partial^2 D^C_{pq}}{\partial\lambda_i\partial\lambda_j}\right).\end{aligned} \tag{A9}$$

[1] K. Burke, "Perspective on density functional theory," J. Chem. Phys. **136**, 150901 (2012).
[2] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Insights into current limitations of density functional theory," Science **321**, 792–794 (2008).
[3] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Challenges for density functional theory," Chem. Rev. **112**, 289–320 (2012).
[4] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, "Multiconfiguration self-consistent field and multireference configuration interaction methods and applications," Chem. Rev. **112**, 108–181 (2012).
[5] K. H. Marti and M. Reiher, "The density matrix renormalization group algorithm in quantum chemistry," Z. Phys. Chem. **224**, 583–599 (2010).
[6] G. K. L. Chan and S. Sharma, "The density matrix renormalization group in quantum chemistry," Annu. Rev. Phys. Chem. **62**, 465–481 (2011).
[7] Y. Kurashige, "Multireference electron correlation methods with density matrix renormalisation group reference functions," Mol. Phys. **112**, 1485–1494 (2014).
[8] S. Knecht, E. D. Hedegård, S. Keller, A. Kovyrshin, Y. Ma, A. Muolo, M. Stein, and C. J. Reiher, "New approaches for *ab initio* calculations of molecules with strong electron correlation," Chimia **70**, 244–251 (2016).
[9] C. Filippi, R. Assaraf, and S. Moroni, "Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo," J. Chem. Phys. **144**, 194105 (2016).

[10]G. Li Manni, S. D. Smart, and A. Alavi, "Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte Carlo within a super-CI framework, with application to challenging metal- porphyrins," J. Chem. Theory Comput. **12**, 1245–1258 (2016).
[11]R. E. Thomas, Q. Sun, A. Alavi, and G. H. Booth, "Stochastic multi-configurational self-consistent field theory," J. Chem. Theory Comput. **11**, 5316–5325 (2015).
[12]J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen, "Determinant based configuration-interaction algorithms for complete and restricted configuration-interaction spaces," J. Chem. Phys. **89**, 2185–2192 (1988).
[13]T. Fleig, J. Olsen, and C. M. Marian, "The generalized active space concept for the relativistic treatment of electron correlation. I. Kramers-restricted two-component configuration interaction," J. Comput. Phys. **114**, 4775 (2001).
[14]J. Thyssen, T. Fleig, and H. J. A. Jensen, "A direct relativistic four-component multi-configuration self-consistent-field method for molecules," J. Comput. Phys. **129**, 034109 (2008).
[15]D. Ma, G. Li Manni, and L. Gagliardi, "The generalized active space concept in multiconfigurational self-consistent field methods," J. Chem. Phys. **135**, 044128 (2011).
[16]K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlev, and K. Wolinski, "Second-order perturbation theory with a CASSCF reference function," J. Phys. Chem. **94**, 5483–5488 (1990).
[17]K. Andersson, P.-Å. Malmqvist, and B. O. Roos, "Second-order perturbation theory with a complete active space self-consistent field reference function," J. Chem. Phys. **96**, 1218–1226 (1992).
[18]P. Å. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, "The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems," J. Chem. Phys. **128**, 204109 (2008).
[19]D. Ma, G. Li Manni, J. Olsen, and L. Gagliardi, "Second-order perturbation theory for generalized active space self-consistent-field wave functions," J. Chem. Theory Comput. **12**, 3208–3213 (2016).
[20]C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu, "Introduction of n-electron valence states for multireference perturbation theory," J. Chem. Phys. **114**, 10252–10264 (2001).
[21]S. Grimme and M. Waletzke, "A combination of Kohn-Sham density functional theory and multi-reference configuration-interaction methods," J. Chem. Phys. **111**, 5645–5655 (1999).
[22]C. M. Marian and N. Gilka, "Performance of the density functional theory/multireference configuration-interaction method on electronic excitation of extended π-systems," J. Chem. Theory Comput. **4**, 1501–1515 (2008).
[23]G. Li Manni, R. K. Carlson, S. Luo, D. Ma, J. Olsen, D. G. Truhlar, and L. Gagliardi, "Multiconfiguration pair-density functional theory," J. Chem. Theory Comput. **10**, 3669–3680 (2014).
[24]L. Gagliardi, D. G. Truhlar, G. Li Manni, R. K. Carlson, C. E. Hoyer, and J. L. Bao, "Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems," Acc. Chem. Res. **50**, 66–73 (2017).
[25]A. Savin and H.-J. Flad, "Density functionals for the Yukawa electron-electron interaction," Int. J. Quantum Chem. **56**, 327–332 (1995).
[26]T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, "Combining long-range configuration interaction with short-range density functionals," Chem. Phys. Lett. **275**, 151 (1997).
[27]E. Fromager, J. Toulouse, and H. J. Aa. Jensen, "On the universality of the long-/short range separation in multiconfigurational density-functional theory," J. Chem. Phys. **126**, 074111 (2007).
[28]K. Sharkas, A. Savin, H. J. A. Jensen, and J. Toulouse, "A multiconfigurational hybrid density-functional theory," J. Chem. Phys. **137**, 044104 (2012).
[29]E. Fromager, S. Knecht, and H. J. Aa. Jensen, "Multi-configuration time-dependent density-functional theory based on range separation," J. Chem. Phys. **138**, 084101 (2013).
[30]E. D. Hedegård, F. Heiden, S. Knecht, E. Fromager, and H. J. Aa. Jensen, "Assessment of charge-transfer excitations with time-dependent, range-separated density functional theory based on long-range MP2 and multi-configurational self-consistent field wave functions," J. Chem. Phys. **139**, 184308 (2013).
[31]E. D. Hedegård, J. M. H. Olsen, S. Knecht, J. Kongsted, and H. J. Aa. Jensen, "Polarizable embedding with a multiconfiguration short-range density functional theory linear response method," J. Chem. Phys. **142**, 114113 (2015).
[32]E. D. Hedegård, S. Knecht, J. S. Kielberg, H. J. Aa. Jensen, and M. Reiher, "Density matrix renormalization group with efficient dynamical electron correlation through range separation," J. Chem. Phys. **142**, 224108 (2015).
[33]E. D. Hedegård, "Assessment of oscillator strengths with multiconfigurational short-range density functional theory for electronic excitations in organic molecules," Mol. Phys. **115**, 26 (2016).
[34]E. D. Hedegård and M. Reiher, "Polarizable embedding density matrix renormalization group," J. Chem. Theory Comput. **12**, 4242–4253 (2016).
[35]M. Hubert, E. D. Hedegård, and H. J. A. Jensen, "Investigation of multiconfigurational short-range density functional theory for electronic excitations in organic molecules," J. Chem. Theory Comput. **12**, 2203–2213 (2016).
[36]M. Hubert, H. J. A. Jensen, and E. D. Hedegård, "Excitation spectra of nucleobases with multiconfigurational density functional theory," J. Phys. Chem. A **120**, 36–43 (2016).
[37]A. Savin, "On degeneracy, near degeneracy and density functional theory," in *Recent Developments of Modern Density Functional Theory*, edited by J. M. Seminario (Elsevier, Amsterdam, 1996), pp. 327–357.
[38]J. Toulouse, A. Savin, and H.-J. Flad, "Short-range exchange-correlation energy of a uniform electron gas with modified electron–electron interaction," Int. J. Quantum Chem. **100**, 1047–1056 (2004).
[39]J. K. Pedersen, "Description of correlation and relativistic effects in calculations of molecular properties," Ph.D. thesis, University of Southern Denmark, 2004.
[40]J. G. Ángyán, I. C. Gerber, A. Savin, and J. Toulouse, "Van der Waals forces in density functional theory: Perturbational long-range electron-interaction corrections," Phys. Rev. A **72**, 012510 (2005).
[41]E. Fromager and H. J. Aa. Jensen, "Self-consistent many-body perturbation theory in range-separated density-functional theory: A one-electron reduced-density-matrix-based formulation," Phys. Rev. A **78**, 022504 (2008).
[42]J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Ángyán, "Adiabatic-connection fluctuation-dissipation density-functional theory based on range separation," Phys. Rev. Lett. **102**, 096404 (2009).
[43]B. G. Janesko, T. M. Henderson, and G. E. Scuseria, J. Chem. Phys. **130**, 081105 (2009).
[44]E. Goll, H.-J. Werner, and H. Stoll, "A short-range gradient-corrected density functional in long-range coupled-cluster calculations for rare gas dimers," Phys. Chem. Chem. Phys. **7**, 3917–3923 (2005).
[45]E. Fromager, R. Cimiraglia, and H. J. Aa. Jensen, "Merging multireference perturbation and density-functional theories by means of range separation: Potential curves for Be_2, Mg_2, and Ca_2," Phys. Rev. A **81**, 024502 (2010).
[46]See http://daltonprogram.org for development version, DALTON, A molecular electronic-structure program, Release Dalton, 2016.
[47]K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansík, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Sałek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. Ågren, "The Dalton quantum chemistry program system," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 269–284 (2013).
[48]S. Paziani, S. Moroni, P. Gori-Giorgi, and G. B. Bachelet, "Local-spin-density functional for multideterminant density functional theory," Phys. Rev. B **73**, 155111 (2006).
[49]E. Goll, H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, and A. Savin, "A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: Application to alkali-metal rare-gas dimers," Chem. Phys. **329**, 276–282 (2006).
[50]C. Schweitzer and R. Schmidt, "Physical mechanisms of generation and deactivation of singlet oxygen," Chem. Rev. **103**, 1685–1757 (2003).
[51]R. Klotz, C. M. Marian, S. D. Peyerimhoff, B. A. Hess, and R. J. Buenker, "Calculation of spin-forbidden radiative transitions using correlated wavefunctions: Lifetimes of b $^1\Sigma^+$, a$^1\Delta$ states in O_2, S_2 and SO," Chem. Phys. **89**, 223–236 (1984).

[52]O. B. Gadzhiev, S. K. Ignatov, M. Y. Kulikov, A. M. Feigin, A. G. Razuvaev, P. G. Sennikov, and O. Schrems, "Structure, energy, and vibrational frequencies of oxygen allotropes on ($n \leq 6$) in the covalently bound and van der Waals forms: *Ab initio* study at the CCSD(T) level," J. Chem. Theory Comput. **9**, 247–262 (2013).
[53]A. Zen, B. L. Trout, and L. Guidoni, "Properties of reactive oxygen species by quantum Monte Carlo properties of reactive oxygen species by quantum Monte Carlo," J. Chem. Phys. **141**, 014305 (2014).
[54]A. Ghosh and P. R. Taylor, "High-level *ab initio* calculations on the energetics of low-lying spin states of biologically relevant transition metal complexes: A first progress report," Curr. Opin. Chem. Biol. **7**, 113–124 (2003).
[55]J. Toulouse, F. Colonna, and A. Savin, "Long-range–short-range separation of the electron-electron interaction in density-functional theory," Phys. Rev. A **70**, 062505 (2004).
[56]T. Saue and T. Helgaker, "Four-component relativistic Kohn–Sham theory," J. Comput. Chem. **23**, 814–823 (2002).
[57]P. Salek, O. Vahtras, T. Helgaker, and H. Ågren, "Density-functional theory of linear and nonlinear time-dependent molecular properties," J. Chem. Phys. **117**, 9630–9645 (2002).
[58]H. J. Aa. Jensen and P. Jørgensen, "A direct approach to second order MCSCF calculations using a norm-extended optimization scheme," J. Chem. Phys. **80**, 1204–1214 (1984).
[59]H. J. Aa. Jensen and H. Ågren, "MCSCF optimization using the direct, restricted step second-order norm-extended optimization method," Chem. Phys. Lett. **110**, 140–144 (1984).
[60]H. J. Aa. Jensen and H. Ågren, "A direct, restricted-step, second order MCSCF program for large scale *ab initio* calculations," Chem. Phys. **104**, 229–250 (1986).
[61]H. J. Aa. Jensen, P. Jørgensen, and H. Ågren, "Efficient optimization of large scale MCSCF wave functions with a restricted step algorithm," J. Chem. Phys. **87**, 451–466 (1987).
[62]H. J. Aa. Jensen, "Electron correlation in molecules using direct second order MCSCF," in *Relativistic and Electronic Correlation in Molecules and Solids*, edited by G. L. Malli (Plenum Press, 1994), pp. 179–206.
[63]P. E. M. Siegbahn, "A new direct CI method for large CI expansions in a small orbital space," Chem. Phys. Lett. **109**, 417–423 (1984).
[64]T. H. Dunning, Jr., "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007–1023 (1989).
[65]T. H. Dunning, Jr., "Gaussian basis functions for use in molecular calculations. I. Contraction of (9s5p) atomic basis sets for the first-row atoms," J. Chem. Phys. **53**, 2823–2833 (1970).
[66]M. Radoń, G. Katarzyna, J. Szklarzewicz, and E. Broclawik, "Spin-state energetics of Fe(III) and Ru(III) aqua complexes: Accurate *ab initio* calculations and evidence for huge solvation effects," J. Chem. Theory Comput. **12**, 1592–1605 (2016).
[67]F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy," Phys. Chem. Chem. Phys. **7**, 3297–3305 (2005).
[68]J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**, 3865–3868 (1996).
[69]J. Toulouse, F. Colonna, and A. Savin, "Short-range exchange and correlation energy density functionals: Beyond the local-density approximation," J. Chem. Phys. **122**, 014110 (2005).
[70]K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules* (Springer, Boston, 1979).
[71]J. P. Perdew, A. Savin, and K. Burke, "Escaping the symmetry dilemma through a pair-density interpretation of spin-density functional theory," Phys. Rev. A **51**, 4531–4541 (1995).
[72]C. K. Jørgensen, "Studies of absorption spectra IV. Some new transition group bands of low intensity," Acta Chem. Scand. **8**, 1502–1512 (1954).
[73]Y. Yang, M. A. Ratner, and G. C. Schatz, "Multireference *ab initio* study of ligand field d-d transitions in octahedral transition-metal oxide clusters," J. Phys. Chem. C **118**, 29196–29208 (2014).
[74]F. Neese, T. Petrenko, D. Ganyushin, and G. Olbrich, "Advanced aspects of *ab initio* theoretical optical spectroscopy of transition metal complexes: Multiplets, spin-orbit coupling and resonance Raman intensities," Coord. Chem. Rev. **251**, 288–327 (2007).
[75]D. Harris, G. H. Loew, and A. Komornicki, "Structure and relative spin-state energetics of $[Fe(H_2O)_6]^{3+}$: A comparison of UHF, Møller-Plesset, nonlocal DFT, and semiempircal INDO/S calculations," J. Phys. Chem. A **101**, 3959–3965 (1997).
[76]T. F. Hughes and R. A. Friesner, "Correcting systematic errors in DFT spin-splitting energetics for transition metal complexes," J. Chem. Theory Comput. **7**, 19–32 (2011).
[77]Y. Yang, M. A. Ratner, and G. C. Schatz, "Computational modeling of octahedral iron oxide clusters: Hexaaquairon(III) and its dimers," J. Phys. Chem. C **117**, 21706–21717 (2013).
[78]W. P. Anderson, W. D. Edwards, and M. C. Zerner, "Calculated spectra of hydrated ions of the first transition-metal series," Inorg. Chem. **25**, 2728–2732 (1986).
[79]E. Goll, M. Ernst, F. Moegle-Hofacker, and H. Stoll, "Development and assessment of a short-range meta-GGA functional," J. Chem. Phys. **130**, 234112 (2009).