Self-Consistent Field algorithms in Restricted Open-Shell Hartree-Fock

Robert Benda, Laurent Vidal, Emmanuel Giner, Eric Cancès

Self-Consistent Field algorithms in Restricted Open-Shell Hatree-Fock

Robert Benda[1,2], Laurent Vidal[1], Emmanuel Giner[3], and Eric Cancès[1,*]

[1]CERMICS, Ecole des Ponts and Inria Paris, 6 & 8 avenue Blaise Pascal, 77455 Marne-la-Vallée, France
[2]LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau, France
[3]Laboratoire de Chimie Théorique (UMR 7616), Sorbonne Université, CNRS, Paris, France
[*]Corresponding author : eric.cances@enpc.fr

Abstract

In this chapter, we propose a simple geometrical derivation of the restricted open-shell Hartree-Fock (ROHF) equations in the density matrix and molecular orbitals formalism. We then introduce a new, parameter-free, basic fixed-point method to solve these equations, that, in contrast with existing self-consistent field (SCF) schemes, is not based on the introduction of a non-physical, parameter-dependent, composite Hamiltonian. We also extend the Optimal Damping Algorithm to the ROHF framework. We finally present numerical results on challenging systems (complexes with transition metals) demonstrating the performance of the new algorithms we propose.

1 Introduction

The ultimate goal of computational chemistry is to propose reliable theoretical tools to describe the chemical properties of any molecular system. The initial step of such a task is always the accurate description of the ground state electronic structure of the system, for which there exist essentially two flavors of approaches: the wave function theory (WFT) and density functional theory (DFT). Although DFT remains certainly the most used theoretical tool for closed-shell systems because of its advantageous ratio between the computational cost and the accuracy of the results, the usual semi-local approximations used in DFT are known to suffer from several issues when open-shell systems need to be considered. For instance, the self-interaction error in open-shell systems is responsible for the over delocalization of electrons in transition metal complexes and has impacts on several chemical properties such as the electronic paramagnetic spectrum, ligand-field excitations or spin-gaps [38, 43, 1, 25, 5]. One major issue in DFT is that there is no systematic way to improve the results, which leads to an inflation of different flavors of approximated functionals tailored for a specific class of systems and/or properties [47]. The situation of WFT is somehow opposite as there exists many ways of systematically refine the results starting from a mean-field description although it comes to the price of a rapidly growing computational cost. Nevertheless, as remarkable progresses have been obtained in the reduction of the computational cost of correlated WFT methods (see for instance Ref [29] and references therein), the latter appear more and more as actual computational tools for the treatment of open-shell systems. Even though WFT-based correlated methods are in active development, they all start with a mean-field Hartree Fock (HF) calculation for which there are many convergence problems in the context of open-shell systems. Therefore, improving the

reliability of the HF algorithms becomes an important point in order to popularize the correlated WFT methods.

There exists several avatars of the Hartree-Fock method. The most commonly used are the restricted and unrestricted Hartree-Fock methods (RHF and UHF, respectively), which differ by the constraint imposed in the RHF method to have an unique set of spatial orbitals for both up and down spins. For open-shell systems, the constraint of having the same spatial orbitals for the two spins has an important consequence: while the ROHF Slater determinant is an eigenfunction of the $\hat{\mathcal{S}}^2$ operator, the UHF Slater determinant suffers from spin contamination [45]. The latter has a big impact in the post-HF calculations as the correlated wave function built upon a spin-contaminated Slater determinant needs to restore the correct spin symmetry using high-order particle-hole excitations [9, 45, 16]. Moreover, the correlated methods using unrestricted orbitals necessary deal with several types of two-electron integrals corresponding to the interaction between electrons of different spins, which also induces several complications in the code structure and memory.

From the mathematical point of view, Hartree-Fock methods give rise to constrained optimization problems, whose first-order optimality conditions are the Hartree-Fock equations. As usual in optimization theory, numerical solutions can be obtained either by solving the Hartree-Fock equations by a fixed-point (self-consistent field - SCF) algorithm, or by a direct minimization of the Hartree-Fock energy functional [16, 48, 13].

Many algorithms have been developed for the RHF and UHF frameworks in the past 70 years. Roothaan's [36], level-shifting [39], and DIIS algorithms [33, 34, 21, 35, 15] belong to the class of SCF algorithms. Direct minimization approaches are adopted in *e.g.* Bacskay's quadradic convergent algorithm [6], trust-region methods [44] and geometric direct minimization (GDM) methods [16, 48]. Let us also mention the second-order SCF (SOSCF) algorithm [14, 31], and the DIIS-GDM [16, 48], which combine features from both SCF and direct minimization methods.The optimal damping algorithm (ODA) [11] and the EDIIS algorithm [26] solve a relaxed version of the Hartree-Fock optimization problem, whose solutions always coincide with those of the original Hartree-Fock problem for UHF, as well as for the less popular General Hartree-Fock method (GHF) in which each spin-orbital is allowed to have both a spin-up and a spin-down component. For RHF, ODA and EDIIS most often converge to solutions to the RHF problem, but may occasionally converge to one-body density matrices with fractional occupation numbers, which do not correspond to Hartree-Fock states. A robust and efficient method to solve the RHF and UHF problems (which always works for UHF and most of the time for RHF) is to use EDIIS in the first iterations and switch to DIIS to accelerate convergence when the iterates are close enough to the solution [26]. All the above algorithms are relatively well-understood from a mathematical point of view [12]. Roughly speaking, computing RHF and UHF ground states for small and medium-size chemical systems is no longer an issue.

The situation is radically different for ROHF, where existing SCF algorithms fail to converge in many cases, notably for radicals and molecular systems containing transition metals.

In this article, we investigate the SCF algorithms for ROHF. We focus on maximum spin states in order to simplify the presentation, but our approach is valid for any spin state (see Remark 2.1). In Section 2, we recall the mathematical structure of the ROHF ground state problem in both the density matrix and molecular orbital formalisms. In particular, we point out that the ROHF minimization space has the geometry of a flag manifold, a structure that has been described in the mathematical literature (see e.g. [27, 50]). Using this formalism, we derive from a geometric perspective the first-order optimality conditions for the ROHF problem, the ROHF equations.

In contrast with the RHF and UHF settings, the ROHF equations cannot be *naturally* formulated as a nonlinear eigenvalue problem. As a consequence, the simple SCF Roothaan scheme for RHF, *"assemble the Fock matrix for the current iterate, diagonalize it, build the next iterate using*

the Aufbau principle, that is by selecting the lowest energy orbitals", cannot be straightforwardly extended to the ROHF setting. All the existing SCF algorithms we are aware of twist the ROHF equations using coupling operators to transform them into a nonlinear eigenvalue problem. They are based on the construction of a composite, non-physical, effective Hamiltonian obtained by linear combinations of sub-blocks of the Fock matrices F_d and F_s respectively associated to the doubly and singly ROHF orbitals (also refereed to as *internal* and *active* orbitals). These combinations involve six real coefficients A_{tt}, and B_{tt} with t equal to d (doubly occupied), s (singly occupied), or v (virtual), the choice of which characterizes the SCF scheme. For instance, these six coefficients are all equal to $1/2$ in the Guest and Saunders algorithm [20], but are different and depend on the spin state in the Canonical-I and Canonical-II algorithms introduced by Plakhutin and Davidson [32]. From the physical point of view, the choice of A_{tt} and B_{tt} coefficients essentially tries to maintain the *Aufbau* principle in order to avoid numerical instabilities of the SCF algorithm induced by swapping of the singly occupied orbital with doubly occupied or virtual orbitals. It is important to stress that, because of the mathematical restriction imposed by the ROHF Slater determinant, the *Aufbau* principle, inspired by the Koopman theorem, is not guaranteed, and therefore a choice of A_{tt} and B_{tt} which might work for a given system might break down for another, as illustrated for instance in the numerical results reported here (see Section 4.2).

In Section 3, we present a new SCF scheme, which better respects the essence of the ROHF equations and which is parameter-free. We then briefly describe how the DIIS acceleration algorithms write on the flag manifold of ROHF states. In Section 3.2.1, we extend the ODA to the ROHF setting. In Section 4, we compare the performance of the new algorithms introduced in this article to the state-of-the-art SCF algorithms for some challenging chemical systems, such as organic ligands chelating – or simply interacting with – transition metals. Although computationally demanding in their current state, our new algorithms showcase robust convergence properties, and give new perspective on the design of black-box SCF algorithms for open-shell systems.

2 The ROHF optimization problem

In this section, we first present the ROHF model in density matrices (DM) and molecular orbitals (MO) formalisms (without virtual orbitals). We then introduce the manifold of ROHF states. This manifold has a rich geometrical structure, known as a flag manifold. Although they are equivalent, each formalism DM or MO produces a specific discretization of the flag manifold of ROHF states, the ROHF energy gradient and optimality conditions, each one providing some insight on the ROHF problem.

2.1 The ROHF model

In ROHF theory, trial wavefunctions Ψ are not, in general, single Slater determinants, but configuration state functions (CSFs) [32, 23]. The latter are eigenfunctions of the spin operators $\hat{S}^2$ and $\hat{S}_z$ and of the number operators $\widehat{n}_i = a_{i\uparrow}^\dagger a_{i\uparrow} + a_{i\downarrow}^\dagger a_{i\downarrow}$, for a given orthonormal basis of orbitals $(\varphi_1, \varphi_2, \cdots)$ of $L^2(\mathbb{R}^3;\mathbb{C})$:

$$\hat{S}^2\Psi = s(s+1)\Psi, \quad \hat{S}_z\Psi = m_s\Psi, \quad \widehat{n}_i\Psi = n_i\Psi,$$

for given $s \in \frac{1}{2}\mathbb{N}$, $m_s \in \{-s, -s+1, \cdots, s-1, s\}$, and $n_i \in \{0, 1, 2\}$. Up to reordering the orbitals, we can assume that

$$\begin{aligned} n_i &= 2 \quad \forall i = 1, \cdots, N_d, \\ n_i &= 1 \quad \forall i = N_d+1, \cdots, N_d+N_s, \\ n_i &= 0 \quad \forall i > N_d+N_s. \end{aligned}$$

Then, Ψ is a finite sum of Slater determinants, each of them made of the N_d doubly occupied orbitals $\varphi_1, \cdots, \varphi_{N_d}$ and N_s spin-orbitals of the form $\varphi_{N_d+1} \otimes \eta_1, \cdots, \varphi_{N_d+N_s} \otimes \eta_{N_s}$, the function η_j being equal to either α (spin-up) or β (spin-down). The numbers N_d, N_s, N (number of electrons in the system), s, and $m_{\rm s}$ are such that

$$2N_d + N_s = N, \quad |m_{\rm s}| \leq s \leq \frac{1}{2}N_s.$$

We also denote by $N_o := N_d + N_s$ the number of (singly or doubly) occupied orbitals.

For maximum spin states ($s = \frac{1}{2}N_s$) and maximum m_s value ($m_{\rm s} = s$), ROHF trial wavefunctions are single Slater determinants built with N_d doubly occupied orbitals $\varphi_1, \cdots, \varphi_{N_d}$ and N_s spin-up-orbitals $\varphi_{N_d+1} \otimes \alpha, \cdots, \varphi_{N_o} \otimes \alpha$, where the φ_i's satisfy $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ for all $1 \leq i, j \leq N_o$. The electronic Hamiltonian

$$H_N = -\frac{1}{2}\sum_{i=1}^{N} \Delta_{\mathbf{r}_i} + \sum_{i=1}^{N} V_{\rm nuc}(\mathbf{r}_i) + \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

being real-valued in the absence of external magnetic field and spin-orbit coupling, we can assume without loss of generality that the orbitals φ_i are real-valued. In order to obtain a computationally tractable model, the φ_i's are expanded in a finite basis set $\mathcal{X} := (\chi_1, \cdots, \chi_{N_b})$ of real-valued functions of the space variable:

$$\varphi_i(\mathbf{r}) = \sum_{\mu=1}^{N_b} [C_o]_{\mu i} \chi_\mu(\mathbf{r}).$$

In practice, the χ_μ's are non-orthogonal atomic orbitals (AO). In order to simplify the presentation, we will however assume here that the basis $\mathcal{X}$ is orthonormal, or equivalently that the overlap matrix is the identity matrix:

$$S_{\mu\nu} := \int_{\mathbb{R}^3} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})\, d\mathbf{r} = \delta_{\mu\nu}.$$

Let us emphasize that we make this simplification for pedagogical purposes only; extending our arguments to non-orthogonal basis sets is a simple exercise. In that setting, the orthonormality constraints on the orbitals imply that C_o is a rectangular orthogonal matrix; in other words, a point of the Stiefel manifold

$$C_o \in \mathrm{St}(N_o; \mathbb{R}^{N_b}) := \{ C_o \in \mathbb{R}^{N_b \times N_o} \text{ s.t. } C_o^T C_o = I_{N_o} \} \tag{2.1}$$

where I_{N_o} denotes the identity matrix of rank N_o. In the following, it will be helpful to decompose C_o as two orthogonal matrices

$$C_o = (C_d | C_s) \quad \text{with} \quad C_d \in \mathbb{R}^{N_b \times N_d} \quad \text{and} \quad C_s \in \mathbb{R}^{N_b \times N_s} \tag{2.2}$$

corresponding to the coefficients of the doubly and singly occupied orbitals respectively.

From C_o, one can construct the density matrices (DM) P_d and P_s

$$P_d := C_d C_d^T \quad \text{and} \quad P_s := C_s C_s^T. \tag{2.3}$$

The matrices P_d and P_s are the basis representations of the orthogonal projectors on the spaces spanned by the doubly and singly occupied orbitals respectively. Recall that a square matrix P is an orthogonal projector if $P^2 = P = P^T$, and that its rank is the integer $\mathrm{tr}(P)$. These matrices represent the one-body density matrices (projectors)

$$\gamma_d = \sum_{i=1}^{N_d} |\varphi_i\rangle\langle\varphi_i| \quad \text{and} \quad \gamma_s = \sum_{i=N_d+1}^{N_o} |\varphi_i\rangle\langle\varphi_i| \tag{2.4}$$

in the basis set $\mathcal{X}$:

$$\gamma_d = \sum_{\mu,\nu=1}^{N_b} [P_d]_{\mu\nu} |\chi_\mu\rangle\langle\chi_\nu| \quad \text{and} \quad \gamma_s = \sum_{\mu,\nu=1}^{N_b} [P_s]_{\mu\nu} |\chi_\mu\rangle\langle\chi_\nu|.$$

We have the following equivalences:

$$\langle\varphi_i|\varphi_j\rangle = \delta_{ij} \text{ for all } 1 \le i,j \le N_o \quad \Leftrightarrow \quad C_d^T C_d = I_{N_d},\ C_s^T C_s = I_{N_s},\ C_d^T C_s = 0 \tag{2.5}$$

$$\Leftrightarrow \quad \left\{ \begin{array}{l} P_d^2 = P_d = P_d^T,\ \mathrm{tr}(P_d) = N_d, \\ P_s^2 = P_s = P_s^T,\ \mathrm{tr}(P_s) = N_s, \\ P_d P_s = 0. \end{array} \right. \tag{2.6}$$

The maximum spin ROHF wavefunction Ψ generated by orthonormal doubly orbitals $(\varphi_1, \cdots, \varphi_{N_b})$ and singly occupied orbitals $(\varphi_{N_d+1}, \cdots, \varphi_{N_o})$ is completely determined (up to an irrelevant global phase) by the one-body density matrices γ_d and γ_s defined by (2.4). Conversely any pair (γ_d, γ_s) of orthogonal projectors satisfying $\mathrm{tr}(\gamma_d) = N_d$, $\mathrm{tr}(\gamma_s) = N_s$, and $\gamma_d\gamma_s = 0$ gives rise to a unique ROHF wavefunction $\Psi^{\mathrm{ROHF}}_{\gamma_d,\gamma_s}$ of maximal spin (up to a global phase), whose energy is a function of (γ_d, γ_s):

$$\mathcal{E}^{\mathrm{ROHF}}(\gamma_d, \gamma_s) := \langle \Psi^{\mathrm{ROHF}}_{\gamma_d,\gamma_s} | H_N | \Psi^{\mathrm{ROHF}}_{\gamma_d,\gamma_s} \rangle.$$

After discretization in the finite basis set $\mathcal{X}$, the ROHF energy functional becomes a function of the matrices P_d and P_s representing γ_d and γ_s in this basis:

$$E(P_d, P_s) := \mathcal{E}^{\mathrm{ROHF}} \left(\sum_{\mu,\nu=1}^{N_b} [P_d]_{\mu\nu} |\chi_\mu\rangle\langle\chi_\nu|, \sum_{\mu,\nu=1}^{N_b} [P_s]_{\mu\nu} |\chi_\mu\rangle\langle\chi_\nu| \right).$$

Standard algebraic manipulations lead to

$$\begin{aligned} E(P_d, P_s) = {} & \mathrm{tr}\left(h(2P_d + P_s)\right) + \mathrm{tr}\left((2J(P_d) - K(P_d))(P_d + P_s)\right) \\ & + \frac{1}{2}\,\mathrm{tr}\left((J(P_s) - K(P_s))P_s\right), \end{aligned} \tag{2.7}$$

where

$$[h]_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla\chi_\mu(\mathbf{r}) \cdot \nabla\chi_\nu(\mathbf{r})\, d\mathbf{r} + \int_{\mathbb{R}^3} V_{\mathrm{nuc}}(\mathbf{r})\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\, d\mathbf{r},$$

$$[J(P)]_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} (\mu\nu|\kappa\lambda) P_{\kappa\lambda}, \quad [K(P)]_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} (\mu\kappa|\nu\lambda) P_{\kappa\lambda},$$

and

$$(\mu\nu|\kappa\lambda) := \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\chi_\kappa(\mathbf{r}')\chi_\lambda(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}'.$$

In the following, we will use the fact that the matrix $h \in \mathbb{R}^{N_b \times N_b}_{\mathrm{sym}}$ is symmetric, and that the functions $J, K : \mathbb{R}^{N_b \times N_b}_{\mathrm{sym}} \to \mathbb{R}^{N_b \times N_b}_{\mathrm{sym}}$ are linear and such that

$$\mathrm{tr}\big(J(P)P'\big) = \mathrm{tr}\big(J(P')P\big),\ \mathrm{tr}\big(K(P)P'\big) = \mathrm{tr}\big(K(P')P\big) \text{ for all } P, P' \in \mathbb{R}^{N_b \times N_b}_{\mathrm{sym}}. \tag{2.8}$$

Note that the trace of P_d is equal to N_d, the number of doubly-occupied orbitals. The fact that each of these orbitals hosts two electrons is taken into account by the factors 2 in the first two terms of the right-hand side of Eq. (2.7). In view of (2.6), the density matrix (DM) formulation of the ROHF ground state problem in the basis $\mathcal{X}$ reads

$$\mathcal{E}^{\mathrm{ROHF}}_* := \min\{E(P_d, P_s),\ (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}}(N_d, N_s; \mathbb{R}^{N_b})\}, \tag{2.9}$$

where

$$\mathcal{M}_{\mathrm{DM}}(N_d,N_s;\mathbb{R}^{N_b}) := \Big\{(P_d,P_s)\in \mathbb{R}^{N_b\times N_b}_{\mathrm{sym}}\times\mathbb{R}^{N_b\times N_b}_{\mathrm{sym}} \mid P_d^2=P_d,\ P_s^2=P_s,\ P_dP_s=0,\\ \mathrm{tr}(P_d)=N_d,\ \mathrm{tr}(P_s)=N_s\Big\}. \tag{2.10}$$

The set $\mathcal{M}_{\mathrm{DM}}$ is the set of admissible pairs of doubly and singly occupied density matrices, that are the pairs of matrices actually representing a maximum spin ROHF state in the basis $\mathcal{X}$.

Remark 2.1. *The optimization problem* (2.9) *corresponds to the ROHF model for maximum spin states* ($|m_s| = s = \frac{1}{2}N_s$). *For other spin states* ($|m_s| \leq s < \frac{1}{2}N_s$), *the ROHF problem still is of the form* (2.9). *The energy functional* E *has a different expression (due to the Fock exchange term coupling only spin-orbitals having the same spin), but remains a sum of linear and bilinear forms in* (P_d, P_s). *See e.g. ref. [23] for the derivation of the non-maximal spin energy expressions using the genealogical coupling scheme. Note that the algorithms presented in this article, although formulated for maximum spin state case, can therefore be straightforwardly extended to any spin state.*

The ROHF energy in MO formalism can be deduced from (2.3) and (2.7), for all $C_o \in \mathrm{St}(N_o,\mathbb{R}^{N_b})$

$$\mathcal{E}(C_o) = E(C_dC_d^T, C_sC_s^T). \tag{2.11}$$

An important difference between the DM and MO formalisms is that an ROHF state is represented by one and only one point of $(P_d,P_s)\in\mathcal{M}_{\mathrm{DM}}(N_d,N_s;\mathbb{R}^{N_b})$ (more precisely, the manifold of ROHF states is diffeomorphic to $\mathcal{M}_{\mathrm{DM}}(N_d,N_s;\mathbb{R}^{N_b})$), while it is represented by an infinity of points in $\mathrm{St}(N_o;\mathbb{R}^{N_b})$, namely the points in the set

$$\left\{C_o\begin{pmatrix}U_d & 0\\ 0 & U_s\end{pmatrix} = (C_dU_d|C_sU_s),\ \text{where } (U_d,U_s)\in\mathcal{O}_{N_d}\times\mathcal{O}_{N_s}\right\}\subset\mathrm{St}(N_o,\mathbb{R}^{N_b}). \tag{2.12}$$

where we denoted $\mathcal{O}_N = \{U\in\mathbb{R}^{N\times N}\text{ s.t. } U^TU = I_N\}$ the orthogonal group of $N\times N$ matrices.

One way to recover the unicity of representation of ROHF states in MO formalism relies on the abstract notion of quotient sets. We introduce the equivalence relation on $\mathrm{St}(N_o;\mathbb{R}^{N_b})$ defined by

$$C_o\sim C_o' \Leftrightarrow \exists (U_d,U_s)\in\mathcal{O}_{N_d}\times\mathcal{O}_{N_s} \text{ such that } C_o' = C_o\begin{pmatrix}U_d & 0\\ 0 & U_s\end{pmatrix}, \tag{2.13}$$

such that the set (2.12) is an equivalence class for the equivalence relation (2.13). Then the set of all equivalence classes (2.12), defined as the quotient

$$\mathcal{M}_{\mathrm{MO}}(N_d,N_s;\mathbb{R}^{N_b}) := \mathrm{St}(N_o;\mathbb{R}^{N_b})/\sim = \mathrm{St}(N_o;\mathbb{R}^{N_b})/(\mathcal{O}_{N_d}\times\mathcal{O}_{N_s}) \tag{2.14}$$

is diffeomorphic to both $\mathcal{M}_{\mathrm{DM}}(N_d,N_s;\mathbb{R}^{N_b})$ and the set of ROHF states. In particular, a ROHF state is represented by one and only one element of $\mathcal{M}_{\mathrm{MO}}(N_d,N_s;\mathbb{R}^{N_b})$. Let us clarify the meaning of this property. An element of the quotient $\mathcal{M}_{\mathrm{MO}}(N_d,N_s;\mathbb{R}^{N_b})$ is by definition an equivalence class (2.13). It can therefore be represented by some $C_o\in\mathrm{St}(N_o;\mathbb{R}^{N_b})$ or by any $C_o' = C_o\begin{pmatrix}U_d & 0\\ 0 & U_s\end{pmatrix}$, for $(U_d,U_s)\in\mathcal{O}_{N_d}\times\mathcal{O}_{N_s}$. Denoting $[\![C_o]\!]$ the equivalence class containing C_o, ($[\![C_o]\!]\in\mathcal{M}_{\mathrm{MO}}(N_d,N_s;\mathbb{R}^{N_b})$), we have

$$[\![C_o]\!] = [\![C_o\begin{pmatrix}U_d & 0\\ 0 & U_s\end{pmatrix}]\!],\quad \forall (U_d,U_s)\in\mathcal{O}_{N_d}\times\mathcal{O}_{N_s}.$$

In addition $\mathcal{E}(C_o) = \mathcal{E}(C_o\begin{pmatrix}U_d & 0\\ 0 & U_s\end{pmatrix})$ (*i.e.* all $C_o\in\mathrm{St}(N_o;\mathbb{R}^{N_b})$ in the same equivalence class have the same ROHF energy), so that $\mathcal{E}$ can be seen as a function from $\mathcal{M}_{\mathrm{MO}}(N_d,N_s;\mathbb{R}^{N_b})$ to $\mathbb{R}$, also

denoted $\mathcal{E}$ for simplicity. We can therefore write the ROHF minimization problem in MO formalism as

$$\mathcal{E}_*^{\mathrm{ROHF}} := \min\{\mathcal{E}([\![C_o]\!]),\ [\![C_o]\!] \in \mathcal{M}_{\mathrm{MO}}(N_d, N_s; \mathbb{R}^{N_b})\}. \tag{2.15}$$

The quotient nature of $\mathcal{M}_{\mathrm{MO}}(N_d, N_s; \mathbb{R}^{N_b})$ is not a mere theoretical tool, but is crucial to build efficient implementations of optimization algorithms in MO representation. Yet, taking into account this specificity of MO formalism would require to introduce additional mathematical objects, which could obscure the main subject of our discussion. For that reason, we will mainly focus in the following on the DM formalism, for which $\mathcal{M}_{\mathrm{DM}}(N_d, N_s; \mathbb{R}^{N_b})$ can be seen as a simple subset of $\mathbb{R}^{N_b \times N_b}_{\mathrm{sym}} \times \mathbb{R}^{N_b \times N_b}_{\mathrm{sym}}$. Additionally, the ROHF energy functional has a simple form in DM representation, which makes the DM formalism well-suited for methodological developments.

From a mathematical point of view, $\mathcal{M}_{\mathrm{DM}}(N_d, N_s; \mathbb{R}^{N_b})$ and $\mathcal{M}_{\mathrm{MO}}(N_d, N_s; \mathbb{R}^{N_b})$ are smooth (*i.e.* infinitely differentiable, C^∞) compact manifolds. While the DM and MO parametrizations of ROHF states seem quite different, they are in fact two representations of a same geometric object, as we now discuss bellow.

2.2 The manifold of ROHF states

The purpose of this section is to give some insights on the manifolds of ROHF states $\mathcal{M}_{\mathrm{DM}}(N_d, N_s; \mathbb{R}^{N_b})$ and $\mathcal{M}_{\mathrm{MO}}(N_d, N_s; \mathbb{R}^{N_b})$. In order to simplify the notations, we will abbreviate the DM and MO sets as $\mathcal{M}_{\mathrm{DM}}$ and $\mathcal{M}_{\mathrm{MO}}$, and denote by x the points in $\mathcal{M}_{\mathrm{DM}}$ and $[\![y]\!]$ the points in $\mathcal{M}_{\mathrm{MO}}$.

Let us start with a point $x = (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}}$. Since P_d is a rank-N_d orthogonal projector (*i.e.* a symmetric matrix fulfilling $P_d^2 = P_d$ and $\mathrm{Tr}(P_d) = N_d$), it can be diagonalized in an orthonormal basis of $\mathbb{R}^{N_b}$ and its only eigenvalues are 1 (multiplicity N_d) and 0 (multiplicity $N_s + N_v$). Likewise, P_s is a rank-N_s orthogonal projector. In addition, as $P_d P_s = 0$, we also have $P_s P_d = (P_d P_s)^T = 0$, which implies that P_d and P_s commute and can therefore be co-diagonalized in the same orthonormal basis. Introducing the projector

$$P_v := I_{N_b} - P_d - P_s$$

on the virtual space (the space spanned by the virtual orbitals), which satisfies $P_v^2 = P_v = P_v^T$, $\mathrm{tr}(P_v) = N_v$, and $P_d P_v = P_s P_v = 0$, we obtain that there exists a unitary matrix $C \in O_{N_b}$ such that

$$P_d = C\mathcal{I}_d C^T, \quad P_s = C\mathcal{I}_s C^T, \quad P_v = C\mathcal{I}_v C^T, \quad CC^T = I_{N_b}, \tag{2.16}$$

where

$$\mathcal{I}_{N_d} = \begin{pmatrix} I_{N_d} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{I}_{N_s} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_{N_s} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{I}_{N_v} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{N_v} \end{pmatrix}. \tag{2.17}$$

The equations (2.16) and (2.17) are equivalent to finding an orthonormal basis of eigenvectors (which form the unitary matrix C) of the projectors and selecting the ones corresponding to the eigenvalue 1. Decomposing C as $(C_d|C_s|C_v)$ we have

$$P_d = C_d C_d^T, \quad P_s = C_s C_s^T \quad P_v = C_v C_v^T. \tag{2.18}$$

In other words, the set C_d (respectively C_s) is the set of N_d (respectively N_s) natural orbitals associated to the density matrix P_d (respectively P_s). The orbitals in C_v are then the orthogonal complement to C_d and C_s. The equations (2.18) provide a one-to-one correspondence between $(P_d, P_s) \in \mathcal{M}_{\mathrm{DM}}$ and the set of occupied natural orbitals $[\![C_o = (C_d|C_s)]\!] \in \mathcal{M}_{\mathrm{MO}}$.

This relation between MO and DM formalism can be seen in a geometrical setting, by considering the spaces

$$\mathcal{V}_d = \mathrm{Span}(\varphi_i,\, i \in \{1,\dots,N_d\}) \quad \text{and} \quad \mathcal{V}_s = \mathrm{Span}(\varphi_i,\, i \in \{N_d+1,\dots,N_d+N_s\}) \tag{2.19}$$

spanned by the doubly and singly occupied orbitals respectively. In a discretization basis $\mathcal{X}$, the pair of spaces $(\mathcal{V}_d, \mathcal{V}_s)$ can be parametrized by the pair $(P_d, P_s) \in \mathcal{M}_{\mathrm{DM}}$ of respective orthogonal projectors onto $\mathcal{V}_d$ and $\mathcal{V}_s$. It is also parametrized by the coefficients C_d and C_s in the discretization basis $\mathcal{X}$ of an orthonormal basis of $\mathcal{V}_d$ and $\mathcal{V}_s$. Now, all basis sets represented by a matrix in $[\![C_o = (C_d|C_s)]\!]$ span the same spaces. Hence the couple $(\mathcal{V}_d, \mathcal{V}_s)$ is parametrized by a single point $[\![C_o]\!] \in \mathcal{M}_{\mathrm{MO}}$. Because of the orthonormality constraints (2.6), $\mathcal{V}_d$ and $\mathcal{V}_s$ verify

$$\begin{cases} \{0_{L^2(\mathbb{R}^3)}\} \subsetneq \mathcal{V}_d \subsetneq \mathcal{V}_d \oplus \mathcal{V}_s \subsetneq \mathrm{Span}\,(\mathcal{X}), \\ \dim(\mathcal{V}_d) = N_d, \quad \dim(\mathcal{V}_d \oplus \mathcal{V}_s) = N_d + N_s. \end{cases} \tag{2.20}$$

Mathematically, the pair of spaces $(\mathcal{V}_d, \mathcal{V}_d \oplus \mathcal{V}_s)$ with property (2.20) is called a flag with dimensions N_d and $N_d + N_s$. The set of all such pair of spaces has been studied in the mathematical literature (see e.g. [27, Example 21.22]). It is a smooth manifold called a *flag manifold* and denoted $\mathrm{Flag}(N_d, N_d+N_s; \mathbb{R}^{N_b})$.

From the above reasoning there is a one-to-one correspondence between ROHF states and points on $\mathrm{Flag}(N_1, N_1+N_2; \mathbb{R}^{N_b})$. In other words, the DM and MO sets are two discretizations of the flag manifold $\mathrm{Flag}(N_1, N_1+N_2; \mathbb{R}^{N_b})$, which writes as the diffeomorphisms

$$\mathcal{M}_{\mathrm{MO}}(N_1, N_2; \mathbb{R}^{N_b}) \simeq \mathrm{Flag}(N_1, N_1+N_2; \mathbb{R}^{N_b}) \simeq \mathcal{M}_{\mathrm{DM}}(N_1, N_2; \mathbb{R}^{N_b}). \tag{2.21}$$

In order to derive the first-order optimality conditions associated to the minimization problem (2.9) (a.k.a. the ROHF equations in DM formalism) from a simple geometrical argument, we have to identify the space $T_x\mathcal{M}_{\mathrm{DM}}$ to a point $x \in \mathcal{M}_{\mathrm{DM}}$ of the manifold, that is the vector space of velocities $q = (Q_d, Q_s) = \dot{p}(0)$ at $t = 0$ for all paths

$$p : [-1, 1] \ni t \mapsto p(t) \in \mathcal{M}_{\mathrm{DM}}, \quad \text{such that} \quad p(0) = x \tag{2.22}$$

drawn on $\mathcal{M}_{\mathrm{DM}}$ (as shown in Fig. 1). Similarly, the ROHF equations in MO formalism are found by identifying the tangent spaces $T_{[\![y]\!]}\mathcal{M}_{\mathrm{MO}}$.

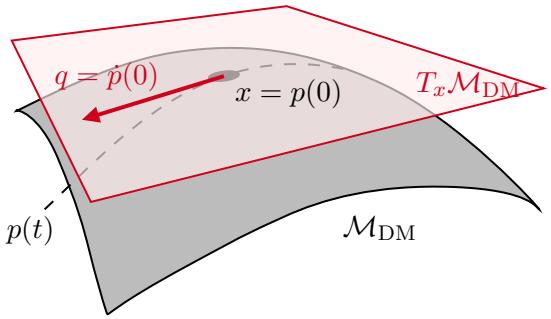

Figure 1: Representation of the tangent space $T_x\mathcal{M}_{\mathrm{DM}}$ at x to the manifold $\mathcal{M}_{\mathrm{DM}}$, and a smooth path $p : [-1, 1] \ni t \mapsto p(t) \in \mathcal{M}_{\mathrm{DM}}$ drawn on $\mathcal{M}_{\mathrm{DM}}$ such that $p(0) = x$ and $\dot{p}(0) = q \in T_x\mathcal{M}_{\mathrm{DM}}$.

Flag manifolds, such as $\mathcal{M}_{\mathrm{DM}}$ and $\mathcal{M}_{\mathrm{MO}}$ have been studied in the context of optimization in the recent work [50], where the authors derive in particular the formulations for the tangent spaces $T_x\mathcal{M}_{\mathrm{DM}}$ and $T_{[\![y]\!]}\mathcal{M}_{\mathrm{MO}}$. To keep this article as self contained as possible, and to make it understandable to readers with limited background in differential geometry, we will adopt in the

following section a pedestrian approach, and re-derive in a few lines the tangent spaces and first order optimality conditions in DM formalism. As mentioned above, details concerning the MO formalism are reported in appendix.

Remark 2.2. *In general, a flag of length d in a vector space $\mathbb{V}$ of dimension N_b is a sequence of subspaces $\{\mathcal{V}_i\}_{1\leqslant i\leqslant d}$ of $\mathbb{V}$ that is strictly increasing for the inclusion. This is to be understood as $\mathcal{V}_1 \subsetneq \cdots \subsetneq \mathcal{V}_d \subsetneq \mathbb{V}$. A standard example of a flag in $\mathbb{V}$ is given by $\{\mathcal{V}_i = \mathrm{Span}(e_1, \cdots, e_i)\}_{1\leqslant i\leqslant N_b}$ where $(e_1, \cdots, e_{N_b})$ is the canonical basis of $\mathbb{V}$. The set of all flags in $\mathbb{V}$ with fixed respective dimensions $\dim(V_i) = n_i$ is also a smooth manifold denoted $\mathrm{Flag}(n_1, \cdots, n_d; \mathbb{V})$ (see e.g. [27, Example 21.22]).*

2.3 First-order optimality conditions

2.3.1 General considerations on optimization in the DM framework

Finding a point $x_* = (P_{d*}, P_{s*})$ in $\mathcal{M}_{\mathrm{DM}}$ which minimizes the energy functional defined in (2.7) requires the definition of the derivative of E with respect to the pair of density matrices $x = (P_d, P_s)$. The ROHF energy functional $E(P_d, P_s)$ is not only defined for density matrices, but for any pair of real-valued symmetric matrix $z = (W_d, W_s) \in \mathbb{R}^{N_b\times N_b}_{\mathrm{sym}} \times \mathbb{R}^{N_b\times N_b}_{\mathrm{sym}}$, which might not be admissible density matrices. Therefore, although the energy gradient $\nabla E(z)$ with respect to $z = (W_d, W_s)$ can be easily computed once a topology, allowing to define the later, has been chosen, imposing $\nabla E(z) = 0$ is not enough to find the optimal ROHF density matrices because of the constraints imposed by the properties of density matrices (see Eq. (2.10)). The reason for this is that the gradient $\nabla E(x)$ has a component outside the manifold $\mathcal{M}_{\mathrm{DM}}$ of density matrices, and following that component of the gradient will necessary lead outside the manifold $\mathcal{M}_{\mathrm{DM}}$. As illustrated in Fig. 2, the correct ROHF condition is therefore to find the point $x_* \in \mathcal{M}_{\mathrm{DM}}$ such that the projection of $\nabla E(x_*)$ onto the tangent space $T_{x_*}\mathcal{M}$ is zero.

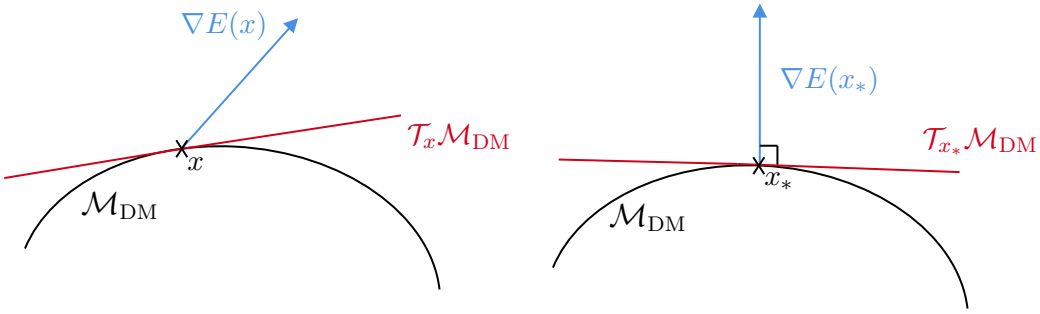

Figure 2: Side view of the manifold $\mathcal{M}_{\mathrm{DM}}$ with tangent space and ambient ROHF energy gradient at (left) arbitrary point x (right) at optimal point x_*. The gradient $\nabla E(x_*)$ is orthogonal to the tangent space $T_{x_*}\mathcal{M}_{\mathrm{DM}}$ (first order optimality conditions).

2.3.2 Characterization of the DM tangent spaces

Let p be a path as in (2.22). We have for all $t \in [-1, 1]$,

$$p(t) \in \mathcal{M}_{\mathrm{DM}} \quad \text{and} \quad p(t) = x + tq + O(t^2) = (P_d + tQ_d + o(t), P_s + tQ_s + o(t)), \tag{2.23}$$

where the $O(.)$ and $o(.)$ notations are relative to the usual Euclidean topology. In other words, the conditions (2.23) are equivalent to defining the tangent space $T_x\mathcal{M}_{\mathrm{DM}}$ to $x = (P_d, P_s)$ as the vector space of pairs of symmetric real matrices $q = (Q_d, Q_s)$ which allow to locally approximate

the manifold of density matrices $\mathcal{M}_{\rm DM}$ by an affine space, as pictorially represented in Fig. 1. The constraints defining the manifold $\mathcal{M}_{\rm DM}$ (see Eq. (2.10)) are equivalent to the following at first order:

$$p_d(t)^2 = p_d(t), \quad \mathrm{tr}(p_d(t)) = N_d \quad \Leftrightarrow \quad P_dQ_d + Q_dP_d = Q_d, \quad \mathrm{tr}(Q_d) = 0, \tag{2.24}$$

$$p_s(t)^2 = p_s(t), \quad \mathrm{tr}(p_s(t)) = N_d \quad \Leftrightarrow \quad P_sQ_s + Q_sP_s = Q_s, \quad \mathrm{tr}(Q_s) = 0, \tag{2.25}$$

$$p_d(t)p_s(t) = 0 \quad \Leftrightarrow \quad P_dQ_s + Q_dP_s = 0. \tag{2.26}$$

In the representation (2.16)-(2.17), the constraints (2.24)-(2.26) are equivalent to

$$Q_d = C \begin{pmatrix} 0 & X & Y \\ X^T & 0 & 0 \\ Y^T & 0 & 0 \end{pmatrix} C^T \quad \text{and} \quad Q_s = C \begin{pmatrix} 0 & -X & 0 \\ -X^T & 0 & Z \\ 0 & Z^T & 0 \end{pmatrix} C^T, \tag{2.27}$$

where $X \in \mathbb{R}^{N_d \times N_s}$, $Y \in \mathbb{R}^{N_d \times N_v}$, $Z \in \mathbb{R}^{N_s \times N_v}$ are generic matrices. It follows that for all $x = (P_d, P_s)$ in $\mathcal{M}_{\rm DM}$:

$$\begin{aligned} T_x\mathcal{M}_{\rm DM} &= \{(Q_d, Q_s) \in \mathcal{V}_{\rm sym} \text{ of the form (2.27)}\} \\ &= \{(Q_d, Q_s) \in \mathcal{V}_{\rm sym} \mid P_dQ_dP_d = P_sQ_dP_s = P_vQ_dP_v = P_sQ_dP_v = 0, \\ &\qquad P_dQ_sP_d = P_sQ_sP_s = P_vQ_sP_v = P_dQ_sP_v = 0,\ P_d(Q_d + Q_s)P_s = 0\}. \end{aligned}$$

2.3.3 ROHF-Brillouin condition in the MO and DM framework

We denote the ambient DM space

$$\mathcal{V}_{\rm DM} = \mathbb{R}^{N_b \times N_b}_{\rm sym} \times \mathbb{R}^{N_b \times N_b}_{\rm sym} \tag{2.28}$$

endowed with the Frobenius-like scalar product

$$\langle (M_1, N_1), (M_2, N_2) \rangle_{\rm DM} := \frac{1}{2} \left(\mathrm{tr}(M_1M_2) + \mathrm{tr}(N_1N_2) \right). \tag{2.29}$$

Thanks to this inner product, the critical points of E on $\mathcal{M}_{\rm DM}$ can be characterized in a simple geometric way (see Fig. 2):

$$x_* \text{ critical point of } E \text{ on } \mathcal{M}_{\rm DM} \quad \Leftrightarrow \quad \nabla E(x_*) \in T_{x_*}\mathcal{M}_{\rm DM}^{\perp}, \tag{2.30}$$

where $\nabla E(x_*)$ is the gradient of E for the inner product $\langle \cdot, \cdot \rangle_{\rm DM}$, and $T_{x_*}\mathcal{M}_{\rm DM}^{\perp}$ the orthogonal subspace to $T_{x_*}\mathcal{M}_{\rm DM}$, still for the inner product $\langle \cdot, \cdot \rangle_{\rm DM}$. The condition of Eq. (2.30) is equivalent to state that, taken at the optimal point x_*, the component of $\nabla E(x_*)$ on the tangent plane $T_{x_*}\mathcal{M}$ is zero. Recall that for any $x \in \mathcal{V}_{\rm DM}$, $\nabla E(x)$ is the vector of $\mathcal{V}_{\rm DM}$ characterized by

$$E(x + \delta x) = E(x) + \langle \nabla E(x), \delta x \rangle_{\rm DM} + o(\delta x),$$

which implies that the gradient depends on the choice of inner product. Also, for any $x \in \mathcal{M}_{\rm DM}$, the vector space $T_x\mathcal{M}_{\rm DM}^{\perp}$ is defined by

$$T_x\mathcal{M}_{\rm DM}^{\perp} = \{q' \in \mathcal{V}_{\rm DM} \mid \forall q \in T_x\mathcal{M},\ \langle q, q' \rangle_{\rm DM} = 0\}.$$

Gradient of E. Let us first detail the computation of $\nabla E(x)$ for any ROHF state $x = (P_d, P_s) \in \mathcal{V}_{\mathrm{DM}}$. Introducing the Fock operators

$$F_d(P_d, P_s) := h + 2J(P_d) + J(P_s) - K(P_d) - \frac{1}{2}K(P_s), \tag{2.31}$$

$$F_s(P_d, P_s) := \frac{1}{2}\left(h + 2J(P_d) + J(P_s) - K(P_d) - K(P_s)\right), \tag{2.32}$$

we have for all $(M_d, M_s) \in \mathcal{V}_{\mathrm{DM}}$

$$\begin{aligned}
E(P_d + M_d, P_s + M_s) &= \mathrm{tr}\left(h(2P_d + 2M_d + P_s + M_s)\right) \\
&\quad + \mathrm{tr}\left((2J(P_d + M_d) - K(P_d + M_d))(P_d + M_d + P_s + M_s)\right) \\
&\quad + \frac{1}{2}\mathrm{tr}\left((J(P_s + M_s) - K(P_s + M_s))(P_s + M_s)\right) \\
&= E(P_d, P_s) + \mathrm{tr}\left(2F_d(P_d, P_s)M_d\right) + \mathrm{tr}\left(2F_s(P_d, P_s)M_s\right) \\
&\quad + \mathrm{tr}\left((2J(M_d) - K(M_d))(M_d + M_s)\right) + \frac{1}{2}\mathrm{tr}\left((J(M_s) - K(M_s))M_s\right) \\
&= E(P_d, P_s) + \langle (4F_d(P_d, P_s), 4F_s(P_d, P_s)), (M_d, M_s)\rangle_{\mathrm{DM}} \\
&\quad + \mathrm{tr}\left((2J(M_d) - K(M_d))(M_d + M_s)\right) + \frac{1}{2}\mathrm{tr}\left((J(M_s) - K(M_s))M_s\right).
\end{aligned}$$

The gradient of E at $x = (P_d, P_s)$ for the inner product $\langle \cdot, \cdot \rangle_{\mathrm{DM}}$ is therefore

$$\nabla E(x) = (4F_d(P_d, P_s), 4F_s(P_d, P_s)) \text{ with } F_d(P_d, P_s) \text{ and } F_s(P_d, P_s) \text{ given by (2.31)-(2.32).} \tag{2.33}$$

Characterization of $T_x\mathcal{M}_{\mathrm{DM}}^{\perp}$. Let $q' = (M_d, M_s) \in \mathcal{V}_{\mathrm{DM}}$. Using the decomposition

$$M_d = U \begin{pmatrix} M_d^{dd} & M_d^{ds} & M_d^{dv} \\ M_d^{sd} & M_d^{ss} & M_d^{sv} \\ M_d^{vd} & M_d^{vs} & M_d^{vv} \end{pmatrix} U^T \quad \text{and} \quad M_s = U \begin{pmatrix} M_s^{dd} & M_s^{ds} & M_s^{dv} \\ M_s^{sd} & M_s^{ss} & M_s^{sv} \\ M_s^{vd} & M_s^{vs} & M_s^{vv} \end{pmatrix} U^T, \tag{2.34}$$

and the fact that M_d and M_s are symmetric matrices, we obtain that for all $q = (Q_d, Q_s) \in T_x\mathcal{M}_{\mathrm{DM}}$ of the form (2.27),

$$\begin{aligned}
\langle q, q'\rangle_{\mathrm{DM}} &= \frac{1}{2}\mathrm{tr}\left(U \begin{pmatrix} 0 & X & Y \\ X^T & 0 & 0 \\ Y^T & 0 & 0 \end{pmatrix} U^T U \begin{pmatrix} M_d^{dd} & M_d^{ds} & M_d^{dv} \\ M_d^{sd} & M_d^{ss} & M_d^{sv} \\ M_d^{vd} & M_d^{vs} & M_d^{vv} \end{pmatrix} U^T\right) \\
&\quad + \frac{1}{2}\mathrm{tr}\left(U \begin{pmatrix} 0 & -X & 0 \\ -X^T & 0 & Z \\ 0 & Z^T & 0 \end{pmatrix} U^T U \begin{pmatrix} M_s^{dd} & M_s^{ds} & M_s^{dv} \\ M_s^{sd} & M_s^{ss} & M_s^{sv} \\ M_s^{vd} & M_s^{vs} & M_s^{vv} \end{pmatrix} U^T\right) \\
\Leftrightarrow \langle q, q'\rangle_{\mathrm{DM}} &= \mathrm{tr}\left(X^T(M_d^{ds} - M_s^{ds})\right) + \mathrm{tr}\left(Y^T M_d^{dv}\right) + \mathrm{tr}\left(Z^T M_s^{sv}\right).
\end{aligned} \tag{2.35}$$

Now, q' belongs to the orthogonal subspace $T_x\mathcal{M}_{\mathrm{DM}}^{\perp}$ if $\langle q, q'\rangle_{\mathrm{DM}} = 0$ for all $q \in T_x\mathcal{M}_{\mathrm{DM}}$. Therefore, according to Eq. (2.35)

$$q' \in T_x\mathcal{M}_{\mathrm{DM}}^{\perp} \quad \Leftrightarrow \quad \left(M_d^{ds} - M_s^{ds} = 0,\ M_d^{dv} = 0,\ M_s^{sv} = 0\right). \tag{2.36}$$

The critical points $x_* = (P_{d*}, P_{s*})$ of E on $\mathcal{M}_{\mathrm{DM}}$ are then characterized by the first-order optimality condition of Eq. (2.30), which according to Eqs. (2.33) and (2.36), leads to

$$(F_{d*} - F_{s*})^{ds} = 0,\ F_{d*}^{dv} = 0,\ F_{s*}^{sv} = 0, \quad \text{with} \quad F_{d*} := F_d(P_{d*}, P_{s*}) \text{ and } F_{s*} := F_s(P_{d*}, P_{s*}).$$

We recover the well-known ROHF optimality conditions (see e.g. [32]), which can also be written as

$$\left\{\begin{array}{l} P_{d*}(F_{d*}-F_{s*})P_{s*}=0, \quad P_{d*}F_{d*}P_{v*}=0, \quad P_{s*}F_{s*}P_{v*}=0, \\ \text{with } F_{d*}:=F_d(P_{d*},P_{s*}) \text{ and } F_{s*}:=F_s(P_{d*},P_{s*}). \end{array}\right. \tag{2.37}$$

We can similarly derive the optimality conditions in the MO representation, by endowing

$$\mathcal{V}_{\rm MO} := \mathbb{R}^{N_b\times N_o}$$

with the Frobenius inner product

$$\langle C_o|C_o'\rangle_{\rm MO} = \mathrm{tr}\left(C_o^TC_o'\right) = \mathrm{tr}\left(C_d^TC_d'\right) + \mathrm{tr}\left(C_s^TC_s'\right). \tag{2.38}$$

This inner product is natural since it reproduces the L^2-inner product. A calculation reported in appendix shows that for all $y=(C_d,C_s)\in\mathcal{V}_{\rm MO}$

$$\nabla\mathcal{E}(y) = (4F_d(C_dC_d^T,C_sC_s^T)C_d, 4F_s(C_dC_d^T,C_sC_s^T)C_s) \tag{2.39}$$

and that $y_*=(C_{d*},C_{s*})\in\mathcal{M}_{\rm MO}$ is a critical point of $\mathcal{E}$ on $\mathcal{M}_{\rm MO}$ if and only if

$$\left\{\begin{array}{l} F_{d*}C_{d*} = C_{d*}(C_{d*}^TF_{d*}C_{d*})\dfrac{1}{2}(C_{s*}(C_{s*}^T(F_{d*}+F_{s*})C_{d*}), \\ \\ F_{s*}C_{s*} = C_{s*}(C_{s*}^TF_{s*}C_{s*}) + \dfrac{1}{2}C_{d*}(C_{d*}^T(F_{s*}+F_{d*})C_{s*}), \\ \\ \text{with } F_{d*}:=F_d(C_{d*}C_{d*}^T,C_{s*}C_{s*}^T) \text{ and } F_{s*}:=F_s(C_{d*}C_{d*}^T,C_{s*}C_{s*}^T). \end{array}\right. \tag{2.40}$$

It can be checked that $C_*=(C_{d*},C_{s*})\in\mathcal{M}_{\rm MO}$ is solution to (2.40) if and only if $(P_{d*},P_{s*})\in\mathcal{M}_{\rm DM}$ is solution to (2.37), where $P_{d_*}:=C_{d*}C_{d*}^T$, $P_{s_*}:=C_{s*}C_{s*}^T$. An important implication of Eqs. (2.40) is that, unlike in the RHF and UHF frameworks, the optimal ROHF orbitals in C_{d*} and C_{s*} are not eigenfunctions of the Fock operators F_{d*} and F_{s*}, because of the second term in the right hand side of the first two equations in (2.40). As a consequence, SCF algorithms based on Fock-like operators involve *ad-hoc* effective Hamiltonians for which the *Aufbau* principal is not always satisfied (see for instance Ref. [32]).

3 Self-consistent field (SCF) algorithms

In this section, we first present the various basic SCF iterations proposed in the literature, and introduce a new one, which better respects the mathematical structure of the ROHF equations (2.37) and (2.40). We then discuss the stabilization and acceleration of basic SCF iterations using Anderson-Pulay (DIIS-type) algorithms.

3.1 Basic SCF iterations

The basic SCF algorithm for RHF was introduced by Roothaan [37]. It consists in assembling the Fock matrix for the current iterate (molecular orbitals or density matrix), diagonalize it (we still assume orthonormality of the basis set for simplicity), and select the lowest energy eigenvectors to form the next iterate (*Aufbau* principle). This idea can be straightforwardly extended to the UHF model, but not to the ROHF model since the ROHF equations (2.40) cannot be formulated as a nonlinear eigenvalue problem.

Let $x^{(k)} = (P_d^{(k)}, P_s^{(k)}) \in \mathcal{M}_{\mathrm{DM}}$ be the current iterate and

$$P_d^{(k)} = C^{(k)} \mathcal{I}_d {C^{(k)}}^T, \quad P_s^{(k)} = C^{(k)} \mathcal{I}_s {C^{(k)}}^T, \quad C^{(k)} {C^{(k)}}^T = I_{N_b},$$

with $C^{(k)} = (C_d^{(k)} | C_s^{(k)} | C_v^{(k)}) \in \mathcal{O}(N_b)$ the associated matrix of natural orbitals via (2.16). Let also $F_d^{(k)} := F_d(P_d^{(k)}, P_s^{(k)})$ and $F_s^{(k)} := F_s(P_d^{(k)}, P_s^{(k)})$ be the associated Fock matrices:

$$F_d^{(k)} = C^{(k)} \begin{pmatrix} {F_d^{(k)}}^{dd} & {F_d^{(k)}}^{ds} & {F_d^{(k)}}^{dv} \\ {F_d^{(k)}}^{sd} & {F_d^{(k)}}^{ss} & {F_d^{(k)}}^{sv} \\ {F_d^{(k)}}^{vd} & {F_d^{(k)}}^{vs} & {F_d^{(k)}}^{vv} \end{pmatrix} {C^{(k)}}^T, \quad F_s^{(k)} = C^{(k)} \begin{pmatrix} {F_s^{(k)}}^{dd} & {F_s^{(k)}}^{ds} & {F_s^{(k)}}^{dv} \\ {F_s^{(k)}}^{sd} & {F_s^{(k)}}^{ss} & {F_s^{(k)}}^{sv} \\ {F_s^{(k)}}^{vd} & {F_s^{(k)}}^{vs} & {F_s^{(k)}}^{vv} \end{pmatrix} {C^{(k)}}^T.$$

3.1.1 Standard approaches

The most popular simple SCF for ROHF consists in assembling and diagonalizing a composite effective Hamiltonian of the form

$$H_{A,B}^{(k)} := C^{(k)} \begin{pmatrix} R_{dd}^{(k)} & (F_d^{(k)} - F_s^{(k)})^{ds} & {F_d^{(k)}}^{dv} \\ (F_d^{(k)} - F_s^{(k)})^{sd} & R_{ss}^{(k)} & {F_s^{(k)}}^{sv} \\ {F_d^{(k)}}^{vd} & {F_s^{(k)}}^{vs} & R_{vv}^{(k)} \end{pmatrix} {C^{(k)}}^T, \tag{3.1}$$

where $R_{dd}^{(k)}$, $R_{ss}^{(k)}$, and $R_{vv}^{(k)}$ are symmetric matrices. The matrices $R_{tt}^{(k)}$ are of the form

$$R_{tt}^{(k)} = 2A_{tt} \left(F_s^{(k)}\right)^{tt} + 2B_{tt} \left(F_d^{(k)} - F_s^{(k)}\right)^{tt}, \quad t \in \{d, s, v\},$$

where $A = (A_{dd}, A_{ss}, A_{vv}) \in \mathbb{R}^3$ and $B = (B_{dd}, B_{ss}, B_{vv}) \in \mathbb{R}^3$ are coefficients characterizing the SCF algorithm (see Table I in [32]). For instance, they are all equal to 1/2 in Guest and Saunders algorithm [20], but are different and depend on the spin state in the Canonical-I and Canonical-II algorithms introduced by Plakhutin and Davidson [32]. The next iterate $(P_d^{(k+1)}, P_s^{(k+1)})$ is obtained by filling up first the doubly occupied orbitals, then the singly occupied orbitals, using the *Aufbau* principle. The meta-algorithm for the basic SCF iteration is summarized in the algorithm 1. The

Algorithm 1: Standard SCF iteration for ROHF

Given: $x^{(k)} = (P_d^{(k)}, P_s^{(k)}) \in \mathcal{M}_{\mathrm{DM}}$, $A = (A_{dd}, A_{ss}, A_{vv})$ and $B = (B_{dd}, B_{ss}, B_{vv})$.

1. Assemble $H_{A,B}^{(k)}$ and diagonalize in an orthonormal basis

$$H_{A,B}^{(k)} C_i^{(k+1)} = \varepsilon_i^{(k+1)} C_i^{(k+1)}, \quad (C_i^{(k+1)})^T C_j^{(k+1)} = \delta_{ij}, \quad \varepsilon_1^{(k+1)} \leqslant \cdots \leqslant \varepsilon_{N_b}^{(k+1)}.$$

2. Select the N_o first orbitals via the *Aufbau principle*

$$C_d^{(k+1)} = (C_1^{(k+1)} | \cdots | C_{N_d}^{(k+1)}), \quad C_s^{(k+1)} = (C_{N_d+1}^{(k+1)} | \cdots | C_{N_d+N_s}^{(k+1)}).$$

3. Construct the new iterate via (2.3)

$$P_d^{(k+1)} = C_d^{(k+1)} {C_d^{(k+1)}}^T, \quad P_s^{(k+1)} = C_s^{(k+1)} {C_s^{(k+1)}}^T, \quad x^{(k+1)} = (P_d^{(k+1)}, P_d^{(k+1)}).$$

iterates are uniquely defined provided

$$\varepsilon_{N_d}^{(k+1)} < \varepsilon_{N_d+1}^{(k+1)} \quad \text{and} \quad \varepsilon_{N_o}^{(k+1)} < \varepsilon_{N_o+1}^{(k+1)} \tag{3.2}$$

(energy gaps between doubly and single-occupied orbitals on the one-hand, occupied and virtual orbitals on the other hand). If the conditions (3.2) are not satisfied, iterates are defined by choosing randomly the orbitals among those satisfying the *Aufbau* principle, or by selecting the ones minimizing the ROHF energy functional. The SCF procedure interprets as a fix point method on the function $g_{A,B} : \mathcal{V}_{\rm DM} \longrightarrow \mathcal{M}_{\rm DM}$ defined by

$$g_{A,B}(x^{(k)}) := x^{(k+1)}, \text{ with } x^{(k+1)} = (P_d^{(k+1)}, P_s^{(k+1)}) \text{ as in algorithm 1.} \tag{3.3}$$

The basic SCF iterations (3.3) being extremely unstable (see section 4), they are generally stabilized by *direct inversion of the iterative subspace* (DIIS) schemes [33, 34, 21, 35, 15].

A necessary and sufficient condition for $(P_{d*}, P_{s*}) \in \mathcal{M}_{\rm DM}$ to be a fixed point of $g_{A,B}$ is

$$H_{A,B*}C_{i*} = \varepsilon_{i*}C_{i*}, \quad C_{i*}^T C_{j*} = \delta_{ij}, \quad \varepsilon_{1*} \leqslant \cdots \leqslant \varepsilon_{N_o*}. \tag{3.4}$$

Let $x_* = (P_{d*}, P_{s*})$ be such a fixed point and C_* the associated matrix of natural orbitals via (2.3). Then

$$P_{d*}H_{A,B*}P_{s*} = P_{d*}(F_{d*} - F_{s*})P_{s*} = \sum_{i=Nd+1}^{N_o} P_{d*}H_{A,B*}C_{i*}C_{i*}^T = \sum_{i=N_d+1}^{N_o} \varepsilon_{i*} \underbrace{P_{d*}C_{i*}}_{=0} C_{i*}^T = 0.$$

A similar argument leads to $P_{d*}F_{d*}P_{v*} = 0$ and $P_{s*}F_{s*}P_{v*} = 0$ so that x_* satisfy the optimality conditions (2.37). Conversely, if x_* satisfies (2.37), then $H_{A,B*} = \mathrm{diag}\,(R_{dd}, R_{ss}, R_{vv})$ is bloc diagonal in the orthogonal decomposition $\mathrm{Ran}(P_{d*}) \oplus \mathrm{Ran}(P_{s*}) \oplus \mathrm{Ran}(P_{v*})$ of $\mathbb{R}^{N_b}$. Therefore, we have

$$\begin{cases} H_{A,B*}C_{i*} = \varepsilon_{i*}C_{i*}, \quad C_{i*}^T C_{j*} = \delta_{ij}, \\ P_{d*} = \displaystyle\sum_{i=1}^{N_d} C_{i*}C_{i*}^T, \quad P_{s*} = \sum_{i=N_d+1}^{N_o} C_{i*}C_{i*}^T, \end{cases}$$

for some orthonormal basis $(C_{i*})_{1\leq i\leq N_b}$ of $\mathbb{R}^{N_b}$ diagonalizing $H_{A,B*}$. It follows that a point $x_* \in \mathcal{M}_{\rm DM}$ is a critical point of E if and only if x_* satisfies the conditions (3.4) *except possibly the fact that the doubly-occupied orbitals do not necessarily correspond to the lowest N_d eigenvalues of $H_{A,B*}$, or the singly-occupied orbitals to the next N_s ones*, which is equivalent to saying that the *Aufbau* principle does not need to be satisfied *a priori*. As discussed in [32], there are indeed local minima of the ROHF problem for which the *Aufbau* principle is not satisfied for any of the usual choices of A and B. We are therefore facing a dilemma. Either the *Aufbau* principle can be kept in the definition of the SCF procedure, leading to a simple iterative scheme, which is however unable to find the ROHF ground state in some cases. Or the *Aufbau* principle can be discarded and replaced by a more complicated construction procedure, to be specified.

3.1.2 A new strategy not based on the *Aufbau* principle

A way out of this dilemma is to attack the problem from a different perspective, using another interpretation of the Roothaan scheme in DM formalism: in the RHF setting, the next iterate $P^{(k+1)}$ obtained by an SCF iteration is the point P of the RHF manifold

$$\mathcal{M}_{\rm DM}^{\rm RHF} := \left\{ P \in \mathbb{R}^{N_b \times N_b}_{\rm sym} \mid P^2 = P,\ \mathrm{tr}(P) = N_d \right\}$$

in the direction along which the slope of the function $t \mapsto E^{\rm RHF}(P^{(k)} + t(P - P^{(k)}))$ is minimum [12], *i.e.*

$$P^{(k+1)} \in \operatorname*{argmin}_{P \in \mathcal{M}_{\rm DM}^{\rm RHF}} \left\langle \nabla E^{\rm RHF}(P^{(k)}) | P \right\rangle_{\mathcal{V}_{\rm DM}^{\rm RHF}} = \operatorname*{argmin}_{P \in \mathcal{M}_{\rm DM}^{\rm RHF}} \mathrm{Tr}\big[F^{\rm RHF}(P^{(k)})P\big] \tag{3.5}$$

where $F^{\mathrm{RHF}}(P) = \frac{1}{2}\nabla E^{\mathrm{RHF}}(P)$ is the Fock matrix associated with the density matrix P, and where $\mathcal{V}_{\mathrm{DM}}^{\mathrm{RHF}} = \mathbb{R}_{\mathrm{sym}}^{N_b \times N_b}$ is the ambient vector space for the RHF problem. In (3.5), argmin refers to the set of minimizers of the linear form $P \longrightarrow \left\langle \nabla E^{\mathrm{RHF}}(P^{(k)}), P \right\rangle_{\mathcal{V}_{\mathrm{DM}}^{\mathrm{RHF}}}$ on $\mathcal{M}_{\mathrm{DM}}^{\mathrm{RHF}}$, to which $P^{(k+1)}$ belongs. This set is always non empty, but may contain several elements. Transposing this characterization to the ROHF setting, we can define a new basic SCF scheme on the manifold $\mathcal{M}_{\mathrm{DM}}$: $x^{(k+1)} := (P_d^{(k+1)}, P_s^{(k+1)})$ is the point $x \in \mathcal{M}_{\mathrm{DM}}$ in the direction along which the slope of the function $t \mapsto E(x^{(k)} + t(x - x^{(k)}))$ is minimum. It is therefore obtained from $x^{(k)} = (P_d^{(k)}, P_s^{(k)})$ as

$$x^{(k+1)} \in \operatorname*{argmin}_{x \in \mathcal{M}_{\mathrm{DM}}} \langle \nabla E(x^{(k)}), x \rangle_{\mathcal{V}_{\mathrm{DM}}} = \operatorname*{argmin}_{x=(P_d,P_s) \in \mathcal{M}_{\mathrm{DM}}} \mathrm{tr}\Big(F_d^{(k)} P_d\Big) + \mathrm{tr}\Big(F_s^{(k)} P_s\Big), \tag{3.6}$$

where $F_d^{(k)} := F_d(P_d^{(k)}, P_s^{(k)})$ and $F_s^{(k)} := F_s(P_d^{(k)}, P_s^{(k)})$. This motivates the introduction of the new basic SCF scheme in algorithm 2 . The fixed points (P_{d*}, P_{s*}) of this SCF scheme verifies

Algorithm 2: New SCF iteration on $\mathcal{M}_{\mathrm{DM}}$

Given: $(P_d^{(k)}, P_s^{(k)}) \in \mathcal{M}_{\mathrm{DM}}$.

1. Compute the Fock matrices $F_d^{(k)} = F_d(P_d^{(k)}, P_s^{(k)})$ and $F_s^{(k)} = F_s(P_d^{(k)}, P_s^{(k)})$
2. Choose next iterate $(P_d^{(k+1)}, P_s^{(k+1)})$ in

$$\operatorname{argmin} \left\{ \mathrm{Tr}\left[F_d^{(k)} P_d + F_s^{(k)} P_s \right], \, (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}} \right\}$$

$$\begin{cases} (P_{d*}, P_{s*}) \in \operatorname{argmin} \{ E_*(P_d, P_s), \ (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}} \}, \\ \quad \text{with} \quad E_*(P_d, P_s) = \mathrm{tr}(F_d(P_{d*}, P_{s*}) P_d) + \mathrm{tr}(F_s(P_{d*}, P_{s*}) P_s). \end{cases} \tag{3.7}$$

Again this SCF procedure can be interpreted as a fix-point method on the function

$$g_{\mathrm{new}}(x^{(k)}) := x^{(k+1)}, \qquad \text{with } x^{(k+1)} = (P_d^{(k+1)}, P_s^{(k+1)}) \text{ as in algorithm 2.} \tag{3.8}$$

As E_* is a linear form, its gradient is constant and equal for the inner product $\langle \cdot, \cdot \rangle_{\mathcal{V}_{\mathrm{DM}}}$ to $(4F_{d*}, 4F_{s*})$. Replacing E with E_* in the arguments in Section 2.3.3, we obtain that (3.7) implies (2.37), hence that any fixed point (P_{d*}, P_{s*}) of the function g_{new} is a critical point of E on $\mathcal{M}_{\mathrm{DM}}$.

The inner optimization problem

$$\operatorname{argmin} \left\{ \mathrm{Tr}\left(F_d^{(k)} P_d + F_s^{(k)} P_s \right), \, (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}} \right\} \tag{3.9}$$

on $\mathcal{M}_{\mathrm{DM}}$ solved at each step is easier and much cheaper to solve numerically than the original problem (2.9) since the function $(P_d, P_s) \mapsto \mathrm{Tr}(F_d^{(k)} P_d + F_s^{(k)} P_s)$ is *linear* while the ROHF energy function $E(P_d, P_s)$ is nonlinear (see Eq. (2.7)). In particular, the Coulomb and Fock terms are not recomputed at each iteration. To solve it, we can use a direct minimization algorithm with initial guess in

$$\operatorname{argmin} \left\{ \mathrm{Tr}(H^{(k)} P_d + \frac{1}{2} H^{(k)} P_s), \ (P_d, P_s) \in \mathcal{M}_{\mathrm{DM}} \right\}, \tag{3.10}$$

where $H^{(k)} = F_d^{(k)}$, or $H^{(k)} = H_{A,B}^{(k)}$, with $H_{A,B}^{(k)}$ given by (3.1). The solutions to (3.10) are easily obtained by diagonalizing $H^{(k)}$ and applying the *Aufbau* principle. For $H^{(k)} = H_{A,B}^{(k)}$, the iterate of the new basic SCF scheme (2) is obtained using $g_{A,B}(P_d^{(k)}, P_s^{(k)})$ as initial guess for the minimization

problem (3.9). It is also possible and more efficient in some cases to use, as an initial guess for the minimization problem (3.9), the previous iterate $(P_d^{(k-1)}, P_s^{(k-1)})$. Let us mention however that this approach only provides *local* (non-necessarily global) minima of (3.9). In practice, we choose for $(P_d^{(k+1)}, P_s^{(k+1)})$ the approximation of the local minimum of $(P_d, P_s) \mapsto \mathrm{Tr}(F_d^{(k)} P_d + F_s^{(k)} P_s)$ on $\mathcal{M}_{\mathrm{DM}}$ obtained by a few iterations of a preconditioned steepest-descent algorithm.

3.2 Anderson-Pulay (DIIS-type) acceleration

Anderson-Pulay acceleration (APA) is a terminology recently coined in [15] to gather various acceleration schemes into a general framework, including the *Anderson acceleration* scheme [4] and the DIIS scheme. Anderson-Pulay acceleration methods can be applied to any fixed-point problems of the form

$$\text{find } x_* \in \mathcal{W} \text{ such that } g(x_*) = x_* \tag{3.11}$$

where $g : \mathcal{W} \to \mathcal{M}$ is a C^2 function from an open subset $\mathcal{W}$ of $\mathbb{R}^n$ into a smooth submanifold $\mathcal{M}$ of $\mathbb{R}^n$. In addition to the fix point map g, APA schemes require a residual function $f : \mathcal{W} \to \mathbb{R}^p$ of class C^2 with $p \leqslant n$, such that for any $x \in \mathcal{W}$, $g(x) = x$ if and only if $f(x) = 0$ (the residual vanishes at solutions to the fixed point problem and only at those points). A possible choice is $f(x) = x - g(x)$ (in which case $p = d$), but the performance of the algorithm can usually be dramatically improved by resorting to well suited residual functions. The APA schemes are based on linear combinations of the current iterate with the previous ones, up to a certain depth $0 \leqslant m \leqslant m_{\max}$. As an example, the standard DIIS acceleration scheme writes for a given depth m, and fix-point map g

$$x^{(k+1)} = g(\mathcal{A}_{\mathrm{DIIS}}(x^{(k)}, \dots x^{(k-m)})) \tag{3.12}$$

where the map $\mathcal{A}_{\mathrm{DIIS}}$ is defined as follows. Let $r^{(k)} := f(x^{(k)})$ and define

$$\mathscr{Y}^{(k)} = \left[x^{(k-m+1)} - x^{(k-m)}, \cdots, x^{(k)} - x^{(k-1)}\right], \quad \mathscr{S}^{(k)} = \left[r^{(k-m+1)} - r^{(k-m)}, \cdots, r^{(k)} - r^{(k-1)}\right].$$

Then

$$\mathcal{A}_{\mathrm{DIIS}}(x^{(k)}, \cdots, x^{(k-m)}) := x^{(k)} + r^{(k)} - (\mathscr{Y}^{(k)} + \mathscr{S}^{(k)})\alpha^{(k)}, \tag{3.13}$$

where the coefficients $\alpha^{(k)} \in \mathbb{R}^m$ are solution to the least square problem

$$\alpha^{(k)} \in \operatorname*{argmin}_{\alpha \in \mathbb{R}^m} \left\| r^{(k)} - \mathscr{S}^{(k)} \alpha \right\|_{\mathbb{R}^p}^2.$$

Mathematical studies on the convergence of DIIS algorithms can be found in [35, 13, 15]. The parameter $m_{\max}$ must be chosen large enough (typically $m_{\max} = 10$ or 20 in quantum chemistry packages) to ensure fast convergence, using sufficient information from previous iterations. One of the limitations of DIIS is that iterates with large residuals (far away from the minimizer) are considered as well, whereas they should be discarded. To cure this deficiency, an adaptive depth approach is proposed in [15], which should be investigated.

Choice of g. In order to be applied to SCF iterations, we need an iteration function defined in an open neighborhood $\mathcal{W}$ of $\mathcal{M}_{\mathrm{DM}}$ since the points $\mathcal{A}(x^{(k)}, \dots, x^{(k-m)})$, which are linear combinations of points of $\mathcal{M}_{\mathrm{DM}}$, do not belong to $\mathcal{M}_{\mathrm{DM}}$ in general. We can directly use one of the basic SCF iteration functions $g_{A,B}$ or g_{new} corresponding to the respective algorithms 1 and 2, since they are defined for any point of $\mathcal{V}_{\mathrm{DM}}$.

Choice of f. From (2.37), a natural choice for the residual function is to take for all $x = (P_d, P_s)$

$$f(P_d, P_s) := ((F_d(P_d, P_s) - F_s(P_d, P_s))^{ds},\ (F_d(P_d, P_s))^{dv},\ (F_s(P_d, P_s))^{sv}) \tag{3.14}$$

which is the projection on $T_x\mathcal{M}_{\rm DM}^{\perp}$ of the gradient $\nabla E(x)$. Remark that this is but a geometrical derivation of the standard commutator based residual used e.g. in GAMESS. In DIIS algorithms, the residual function f is only evaluated at points of the manifold $\mathcal{M}_{\rm DM}$, but must have a C^2 extension to $\mathcal{W}$ for local convergence to be mathematically guaranteed [15]. This is obviously the case for the function f defined by (3.14) on $\mathcal{M}_{\rm DM}$.

3.2.1 Relaxed constrained algorithms for ROHF

Relaxed constrained algorithms for the Unrestricted and General Hartree-Fock setting were introduced in [11]. They consist in optimizing the energy functional in the DM formulation on the convex hull of the admissible set. For the UHF and GHF problems, it can be shown that the relaxed constrained problem has the same global minimizers as the original one [10, 13]. The advantage of the relaxed constrained problems is that convex combinations of admissible solutions are admissible solutions as well.

Algorithm 3: ODA iteration for ROHF

Given: current Fock-like matrices $(\widetilde{F}_d^{(k)}, \widetilde{F}_s^{(k)})$

1. Pick $(P_d^{(k+1)}, P_s^{(k+1)}) \in \operatorname{argmin}\left\{\operatorname{Tr}\left(\widetilde{F}_d^{(k)}P_d + \widetilde{F}_s^{(k)}P_s\right),\ (P_d, P_s) \in \mathcal{M}_{\rm DM}\right\}$
2. Compute the Fock matrices $F_d^{(k+1)} := F_d(P_d^{(k+1)}, P_s^{(k+1)})$, $F_s^{(k+1)} := F_s(P_d^{(k+1)}, P_s^{(k+1)})$ and set
$$(\widetilde{P}_d^{(k+1)}, \widetilde{P}_s^{(k+1)}) = (1-t_k)(\widetilde{P}_d^{(k)}, \widetilde{P}_s^{(k)}), +t_k(P_d^{(k+1)}, P_s^{(k+1)})$$
$$(\widetilde{F}_d^{(k+1)}, \widetilde{F}_s^{(k+1)}) = (1-t_k)(\widetilde{F}_d^{(k)}, \widetilde{F}_s^{(k)}), +t_k(F_d^{(k+1)}, F_s^{(k+1)})$$
where t_k is the minimizer of the quadratic function
$$[0,1] \ni t \mapsto E((1-t)(\widetilde{P}_d^{(k)}, \widetilde{P}_s^{(k)}) + t(P_d^{(k+1)}, P_s^{(k+1)})).$$

The simplest relaxed constrained algorithm is the optimal damping algorithm (ODA). It generates two sequences of iterates:

- a sequence $(x^{(k)})$ of points on the admissible manifold $\mathcal{M}_{\rm DM}$;
- a sequence $(\widetilde{x}^{(k)})$ of points in the convex hull of $\mathcal{M}_{\rm DM}$.

The point $\widetilde{x}^{(k+1)}$ is obtained by doing an optimal convex combination of $\widetilde{x}^{(k)}$ and $x^{(k+1)}$:

$$t_k = \operatorname*{argmin}_{t\in[0,1]} E(tx^{(k+1)} + (1-t)\widetilde{x}^{(k)}), \quad \widetilde{x}^{(k+1)} = t_k x^{(k+1)} + (1-t_k)\widetilde{x}^{(k)}.$$

The function $p_k(t) := E(tx^{(k+1)} + (1-t)\widetilde{x}^{(k)})$ is a second degree polynomial and we have

$$p_k(0) = E(\widetilde{x}^{(k)}) \quad \text{and} \quad p_k'(0) = \langle \nabla E(\widetilde{x}^{(k)}), x^{(k+1)} - \widetilde{x}^{(k)} \rangle_{\mathcal{V}_{\rm DM}}.$$

Computing $p_k(1) = E(x^{(k+1)})$, we obtain the value of t_k explicitly. The point $x^{(k+1)}$ is chosen so as to minimize the slope $p_k'(0)$; it is therefore obtained from $\widetilde{x}^{(k)}$ as

$$x^{(k+1)} \in \operatorname*{argmin}_{x\in\mathcal{M}_{\rm DM}} \langle \nabla E(\widetilde{x}^{(k)}), x \rangle_{\mathcal{V}_{\rm DM}} = g_{\rm new}(\widetilde{x}^{(k)}),$$

where $g_{\rm new}$ is defined in (3.8). The ODA is initialized by choosing an initial guess $x^{(0)} = (P_d^{(0)}, P_s^{(0)})$ in $\mathcal{M}_{\rm DM}$, by setting $\widetilde{x}^{(0)} = x^{(0)}$, and by computing $(\widetilde{F}_d^{(0)}, \widetilde{F}_s^{(0)}) = (F_d(P_d^{(0)}, P_s^{(0)}), F_s(P_d^{(0)}, P_s^{(0)}))$. One then performs ODA iteration as written in algorithm 3.

4 Numerical results

4.1 Methodology and summary of the results

We now analyze the performance of the algorithms introduced in this article which are

- the standard SCF (algorithm 1) and new SCF (algorithm 2), with respective fix point map $g_{A,B}$ and g_{new}, endowed with a DIIS acceleration with residual f given by (3.14);
- the ODA scheme as described in algorithm 3.

Convergence behaviors are investigated in two distinct regimes:

- the global convergence regime. The goal here is to reach the vicinity of a minimizer, starting from a bad initial guess obtained in practice by diagonalizing the core Hamiltonian;
- the local convergence regime, when the initial guess is close to a minimizer. We choose in this study the extended Hückel initial guess derived from the Wolfsberg-Helmholtz approximation [49, 24, 3].

Our implementation. The application of the g_{new} map requires to solve the inner optimization problem (3.9). In our implementation, we use the initial guess (3.10) with $H^{(k)} = F_d^{(k)}$. We then apply a maximum of 10 iterations of preconditioned steepest descent on the DM manifold.

For the ODA method, it happens in some cases that the coefficient t_k of the ODA convex combination becomes zero, which results in the algorithm getting stuck on the iterate $x^{(k)}$. In that case, we automatically try a different guess for the inner problem (3.9). Using a guess generated with $g_{A,B}$ and Euler or Guest and Saunders coefficients whenever $t_k = 0$ proved effective in all the cases we encountered.

The new algorithms we introduce, along with the classical SCF schemes, have been implemented in a `Julia` [7] package as a proof of concept. This package is built as an overlay to the `PySCF` [42] python library, which handles the core computations for ROHF (generation of the AO basis and initial MOs, computation of the electronic integrals). Comprehensive details of implementation can be found in our open-source research code `https://github.com/LaurentVidal95/ROHFToolkit`. The best performing algorithms will be added as a plugin within the Quantum Package [40, 19] and made freely available to the community.

Comparison to external code. In order to assert the validity of our code, we compare the performances of our algorithms with the SCF algorithms for ROHF available in `GAMESS` [41]. We have chosen this popular software because all the classical functions $g_{A,B}$ are implemented, as well as the residual (3.14) for DIIS and the SOSCF algorithm. We have also run tests with `PySCF` and `Psi4` [46] (which respectively implement Roothan and Guest and Saunders' $g_{A,B}$). The DIIS residual functions implemented in these codes can be slightly different but all also use commutator-based residual functions *à la* Pulay [34], involving the effective Hamiltonian $H_{A,B}$.

The initial guesses for the SCF problem, the one generated by `GAMESS` and those employed in our implementation (generated by `PySCF`), can differ significantly. Specifically, the extended Hückel guess in `GAMESS` tends to yield energies approximately 1 to 2 Ha above the ground state energy in our test cases, while the `PySCF` Hückel guesses produce initial energies ranging from 20 to 60 Ha above the ground state. To ensure methodological consistency, and facilitate the direct comparison between the two codes, we manually imported the `GAMESS` Hückel guess in our code for the 6-31G basis set [22, 18, 8] for some of our test cases. We observed no qualitative difference for this choice of basis set. Unfortunately, the two quantum chemistry packages employ different conventions in generating atomic basis sets, particularly concerning the number and order of the atomic orbitals,

which makes the systematic import of GAMESS guesses in our code a laborious task. The comparison with GAMESS should therefore only serve as a qualitative evaluation of our implementation.

Global convergence regime. First, the algorithms are tested by starting *very far* from an expected minimum, *i.e.* starting from a core Hamiltonian diagonalization guess, obtained with GAMESS and PySCF respectively. Poor quality guesses do not usually verify the *Aufbau* principle on which the classical SCF methods, built with function $g_{A,B}$ (3.3), rely (as recalled section 3). Numerical results presented in Section 4.3.1 confirm that, unlike the classical SCF methods for ROHF, which mostly fail to converge in this regime (in all the tested cases but the simplest one), our methods built on g_{new}, which are free of *Aufbau* principle requirement, exhibit a strong robustness with respect to the initial guess.

Local convergence regime. As detailed in section 4.3.2, existing methods built on the classical $g_{A,B}$ barely benefit from the use of an extended Hückel guess, which is more commonly used in practice. Only two or four choices of A_{tt} and B_{tt} coefficients, depending on the test case, yield convergence for these so-called $g_{A,B}$-based methods (see Table 4), with the Guest and Saunders choice being the most successful. Our g_{new}–based methods, that are free of the choice of such coefficients, manage to converge in all cases from this starting guess.

Local minima. The respective $g_{A,B}$-based methods, as well as our g_{new}-based methods, converge toward a variety of local minima. The list of all minima have been reported in appendix. Note that the variation in the implementation of basis sets between GAMESS and our code results in a minor difference in energies. A detailed analysis of the encountered local minima, reached from the core guess and from the Hückel guess, would be needed to assess their quality. It appears that in some cases, the local minima found by starting from the core initial guess, are lower in energy than other minima reached from extended Hückel initial guesses. One should elaborate further on this point in another study.

Our best performing method. When focusing on the energy only, the ODA algorithm seems to target a low minima, independently of the initial guess, while being very slow to converge to chemical accuracy. Applying a few iterations of ODA, followed by g_{new}+DIIS to help convergence is a good candidate for an efficient black-box SCF less sensitive to the initial starting point (see Table 5).

Throughout the next sections, qualitative convergence results are tagged with the following convention:

- ☒ non-convergence: the energies of the iterates oscillate above the ground state energy by at least 10^{-2} Ha and the residual does not go to zero. In many cases, the oscillations occur between 1 and 100 Ha above the ground state energy;
- ☒ stagnation or small-amplitude oscillations : the algorithm stalls or the iterates display small-amplitude oscillations while the residual is small but not small enough in the sense that the limit values of the energy are 10^{-4} to 10^{-2} Ha higher than the ground state energy (or another local minimum)
- ☑ convergence to a local minimizer.

4.2 Basic SCF iterations

We first illustrate the limitations of the classical iteration functions $g_{\mathrm{A,B}}$, as defined in (3.3), and the relevance of the new iteration function g_{new} defined in (3.8), by analyzing the behavior of the corresponding basic SCF algorithms $x^{(k+1)} = g(x^{(k)})$ (without any stabilization/acceleration technique) on simple mono-atomic systems: an oxygen atom in the triplet state, Fe^{2+} and Fe^{3+} ions in high-spin configurations (respectively quintet and sextet states).

Recall that the function $g_{A,B}$ is computed by diagonalizing an effective Hamiltonian depending on the input ROHF state and *ad hoc* coefficients A_{tt} and B_{tt}, and constructing the output ROHF state using the *Aufbau* principle (see Section 3). The performance of the basic SCF algorithm $x^{(k+1)} = g_{A,B}(x^{(k)})$ is found to be very sensitive to the choice of the A_{tt} and B_{tt} coefficients; besides, no choice of coefficients provides consistent convergence for the three simple systems. In contrast, the basic fixed-point algorithm built upon the parameter-free iteration function g_{new} has been able to converge for the three systems. The results reported in Table 1 have been obtained with the double-zeta correlation-consistent Dunning's type basis set (cc-pVDZ) [17] and the Hückel initial guess from `PySCF`. Qualitatively similar results have been obtained with the core initial guess and/or other basis sets (*e.g* 6-31G, pc-1).

Method	A_{tt}	B_{tt}	O (triplet)	Fe^{2+} (quintet)	Fe^{3+} (sextet)
Roothan	$(-\frac{1}{2}, \frac{1}{2}, \frac{3}{2})$	$(\frac{3}{2}, \frac{1}{2}, -\frac{1}{2})$	✓(17)		✓(45)
McWeeny and Diercksen	$(\frac{1}{3}, \frac{1}{3}, \frac{2}{3})$	$(\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$	✓(13)		
Davidson	$(\frac{1}{2}, 1, 1)$	$(\frac{1}{2}, 0, 0)$			✓(12)
Guest and Saunders	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	✓(11)		✓(22)
Binkley, Pople and Dobosh	$(\frac{1}{2}, 1, 0)$	$(\frac{1}{2}, 0, 1)$			✓(10)
Faegri and Manne	$(\frac{1}{2}, 1, \frac{1}{2})$	$(\frac{1}{2}, 0, \frac{1}{2})$			✓(11)
Euler equations	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, 0, \frac{1}{2})$	✓(10)		
Canonical-ROHF I	$(\frac{2S+1}{2S}, 1, 1)$	$(-\frac{1}{2S}, 0, 0)$			✓(11)
Canonical-ROHF II	$(0, 0, -\frac{1}{2S})$	$(1, 1, \frac{2S+1}{2S})$	✓(20)		
g_{new} (3.8)	parameter free		✓(10)	✓(21)	✓(12)

Table 1: Convergence of the basic fixed-point algorithm $x^{(k+1)} = g(x^{(k)})$ for the atomic systems O, Fe^{2+}, and Fe^{3+} (cc-pVDZ basis set, `PySCF` Hückel initial guess), for (i) the classical $g_{A,B}$ iteration functions (see Table I in Ref. [32]), and (ii) the g_{new} iteration function (this work). The table follows the conventions detailed in the introduction to Section 4. The number of iterations needed to reach convergence is specified when the algorithm happens to converge (chosen convergence criterion: the energy of the current iterate is at most 10^{-6} Ha above the ROHF ground state).

4.3 Stabilized and accelerated iteration schemes

Table 2 summarizes the benchmark systems considered in this section. They consist of organic molecules bearing aromatic moieties (such as pyridine or porphyrin), interacting with open-shell

metallic ions (see Figure 3). These systems are representative of the complexity of open-shell calculations in quantum chemistry as they contain transition metal ions with high spin in interaction with non trivial aromatic organic ligands [28]. The combination of strong repulsion in the 3d shell of the metals together with the very delocalized character of the π system in these organic ligands can lead to SCF instabilities precisely because, according to the choice of the flavor of effective Hamiltonian used in the g_{AB} function, the *Aufbau* principle is not fulfilled in these systems. We have picked up both systems having space symmetries, such as pyridine–Cu^{2+} (C_s symmetry) and the Porphyrin model–Fe^{2+} (D_{4h} symmetry), and systems with slightly broken symmetry, such as Pyridine–Fe^{n+}. We infer the spin multiplicities $M = 2S + 1$ of these systems (where S is the total spin) from the corresponding spin multiplicities of the metallic ions, following Hund's rule. In some cases, it is actually challenging to determine the spin multiplicity of the ground state (*e.g.* triplet or quintet), such as for the iron–porphyrin model system [28]. We have performed some test calculations on a full Porphyrin–Fe^{2+} system (37 atoms, 269 basis functions for 6-31G), that yielded qualitatively similar results as for the Porphyrin model–Fe^{2+} system. For the sake of brevity, we do not report them here.

Table 2: Benchmark systems used in Section 4.3.

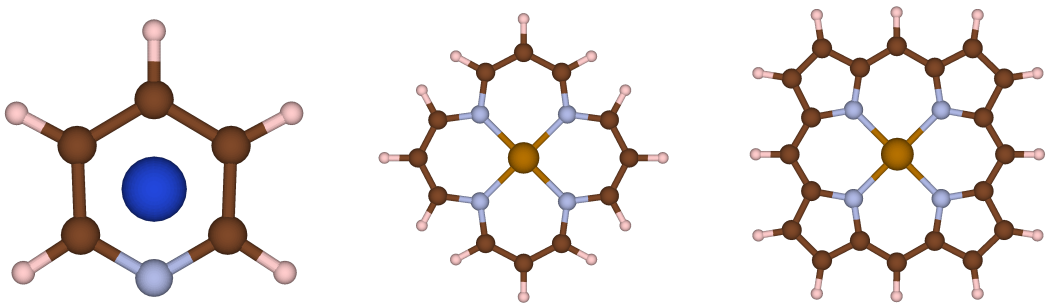

Figure 3: Left: Pyridine - Cu^{2+}. Middle: Porphyrin model – Fe^{2+} taken from [28]. Right: Porphyrin – Fe^{2+}. Figures have been generated with the Vesta software [30].

We have tested several families of basis sets representative of quantum chemistry calculations, *i.e.* the 6-31G and cc-pVDZ basis sets.

4.3.1 Global convergence regime

In this section, we analyze the ability of the various algorithms described in Section 3 to reach the vicinity of a local minimizer from the core initial guess. We consider that this is achieved if the energies of the iterates approach 0.1 Ha from the ROHF ground state energy. We compare the new algorithms proposed in this work with existing algorithms as implemented in GAMESS [41], namely the SOSCF algorithm and the DIIS schemes built from the iteration functions $g_{A,B}$ and residual function f (3.14). The results for the molecular systems in Table 2 in the 6-31G basis set are gathered in Table 3.

Algorithms based on $g_{A,B}$ iteration functions. We observe in the second and third columns of Table 3 that the results of the `GAMESS` implementation of DIIS are close to our DIIS implementation. For several choices of coupling coefficients A_{tt}, B_{tt}, the standard SCF+DIIS method fails to converge, and leads to oscillations. For the Pyridine$-Fe^{2+}$ and Fe^{3+} systems, and (respectively) for the Porphyrin model $-$ Fe^{2+} system, only three (resp. two) specific choices of A_{tt}, B_{tt} coefficients lead

to convergence (notably Guest and Saunders and Roothaan). The results for the Pyridine$-Cu^{2+}$ system (not reported) are qualitatively the same (only Guest and Saunders, Euler, Roothaan and Canonical II choices of A_{tt}, B_{tt} coefficients lead to convergence of GAMESS DIIS or of our DIIS implementation).

Remarkably, forcing DIIS (resp. SOSCF) from the first iterations is needed in `GAMESS`, as the DIIS residual (resp. gradient norm) is initially much higher than the default threshold for DIIS (resp. SOSCF) activation. Let us underline that acceleration methods such as DIIS, are designed to accelerate local convergence (*i.e.* convergence when starting close enough to a local minimum). They are now well-understood mathematically in this setting [15]. In contrast, the fact that DIIS can stabilize SCF iterations starting from core initial guess in some cases (this is not always true) remains unexplained to our knowledge.

The SOSCF second-order method converges whatever the choice of A_{tt}, B_{tt} coefficients (except one, namely Canonical II, for the Pyridine$-Fe^{2+}$ system) from the core guess, although always in more than 200 iterations. Forcing DIIS (resp. SOSCF) from the first iterations is needed in GAMESS, as the DIIS residual (resp. gradient norm) is initially much higher than the default threshold for DIIS (resp. SOSCF) activation.

Algorithms based on the g_{new} iteration function. As shown in the last two columns of Table 3, the DIIS algorithm based on the iteration function g_{new} and the residual function f, as well as the ODA algorithm 3, provide robust schemes for all systems, except for the case of porphyrin model-Fe^{2+} with g_{new}+DIIS. Forcing a restart of the DIIS yields convergence in that case. Note that our current implementation was built as a proof-of-concept. Our method could potentially benefit from a more refined choice of preconditioning for the resolution of the subproblem (3.9), or from an adaptive depth DIIS approach, as introduced in [15], which we defer to future investigations.

For the other cases, the g_{new}+DIIS method is competitive with the converging standard SCF schemes in terms of iterations. The g_{new}+DIIS require more computational time than the $g_{\text{A,B}}$ standard SCFs, since each iteration involves the approximate resolution of the optimization problem (3.9). This is compensated by the absence of parameters in this method, and the convergence across almost all studied cases.

While the ODA method is very effective to reach the attraction basin of a local minimizer, it is very slow to converge to chemical accuracy. As the iterations approach a local minimum, the coefficient t_k of the ODA convex combination consistently equals 1, effectively reducing ODA to a simple SCF with g_{new} map and no DIIS. A good compromise is to transition from ODA to g_{new}+DIIS when sufficiently close to a local minimum (Table 5), mimicking the efficient EDIIS+DIIS method of [26] in the RHF case. This transition can occur when the energy gradient reaches a specified tolerance, or when the ODA coefficient t_k takes the value 1 repeatedly. We chose the first option with threshold 10^{-1} in our implementation. Notably, applying ODA before g_{new}+DIIS seem to allow to target a lower local minimum, as appearing in appendix, Table 6.

	GAMESS [41]		This work		
	$g_{A,B}$ (3.3) - based methods			g_{new} (3.8) - based methods	
A_{tt}, B_{tt} (see Table 1)	SOSCF	DIIS	DIIS	DIIS	ODA
Pyridine–Fe^{2+}					
Guest and Saunders	✓(244;313)	✓(12;37)	✓(9;59)		
Roothaan	✓(212;263)	✓(28;109)	✓(13;145)		
Euler	✓(218;265)	✓(28;95)			
Mc Weeny	✓(204;254)				
Binkley	✓(262;352)			✓(8;55)	✓(4;+1000)
Faegri	✓(235;278)				
Davidson	✓(230;273)				
Canonical I	✓(262;329)				
Canonical II					
Pyridine–Fe^{3+}					
Guest and Saunders	✓(236;290)	✓(16;132)	✓(11;193)		
Roothaan	✓(221;263)	✓(19;72)	✓(17;116)		
Euler	✓(227;277)	✓(41;181)	✓(7;112)		
Mc Weeny	✓(217;273)				
Binkley	✓(216;272)			✓(8;54)	✓(8;+1000)
Faegri	✓(323;374)				
Davidson	✓(259;328)				
Canonical I	✓(246;317)				
Canonical II	✓(236;305)				
Porphyrin model–Fe^{2+}					
Guest and Saunders	✓(202;215)	✓(15;22)	✓(18;26)		
Roothaan	✓(203;219)	✓(21;34)	✓(34;49)		
Euler	✓(202;218)				
Mc Weeny	✓(202;219)				
Binkley	✓(203;213)				✓(10,+1000)
Faegri	✓(203;216)				
Davidson	✓(203;221)				
Canonical I	✓(203;212)				
Canonical II	✓(294;346)				

Table 3: Convergence results starting from core initial guess (6-31G basis set). The table follows the conventions detailed in the introduction to Section 4. The DIIS residual function f is the one defined in (3.14). The DIIS maximum depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). The notation $(n_{\text{approach}}; n_{\text{cv}})$ means that n_{approach} iterations are needed to reach 0.1 Ha accuracy, while n_{cv} iterations are necessary to reach microHartree accuracy.

	GAMESS [41]		This work		
	$g_{A,B}$ (3.3) - based methods			g_{new} (3.8) - based methods	
A_{tt}, B_{tt} (see Table 1)	SOSCF	DIIS	DIIS	DIIS	ODA
Pyridine–Fe^{2+}					
Guest and Saunders	✓(78)	✓(82)	✓(100)	✓(92)	✓(+1000)
Roothaan	✓(83)	✓(255)	✓(212)		
Euler	✓(40)	✓(59)	✓(68)		
McWeeny	✓(42)	✓(105)	✓(271)		
Binkley	✓(106)				
Faegri	✓(106)				
Davidson	✓(87)				
Canonical I	✓(88)				
Canonical II	✓(42)				
Pyridine–Fe^{3+}					
Guest and Saunders	✓(78)	✓(178)	✓(187)	✓(142)	✓(+1000)
Roothaan	✓(88)	✓(185)	✓(139)		
Euler	✓(50)				
McWeeny	✓(88)				
Binkley	✓(93)				
Faegri	✓(92)				
Davidson	✓(94)				
Canonical I	✓(95)				
Canonical II	✓(54)				
Porphyrin model–Fe^{2+}					
Guest and Saunders	✓(22)	✓(17)		✓(25)	✓(+1000)
Roothaan	✓(23)	✓(37)	✓(52)		
Euler	✓(29)	✓(25)	✓(72)		
Mc Weeny	✓(36)	✓(32)	✓(187)		
Binkley	✓(23)				
Faegri	✓(22)				
Davidson	✓(21)				
Canonical I	✓(24)				
Canonical II	✓(29)				

Table 4: Convergence results starting an extended Hückel initial guess (6-31G basis set). The table follows the conventions detailed in the introduction to Section 4. The DIIS residual function f is the one defined in (3.14). The DIIS maximum depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). The number of iterations in parentheses is the one needed to reach microHartree accuracy.

Table 5: Convergence results by starting with ODA iterations and switching to DIIS when the residual norm reaches a tolerance of 10^{-2}. The DIIS depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). The number of iterations needed to reach convergence at microHartree precision is specified in parenthesis

4.3.2 Local convergence

We now compare the different algorithms starting from an extended Hückel initial guess, whose energy is about 1 to 2 Ha above the ground state for our test cases in the GAMESS implementation, and 20 to 60 Ha for PySCF. The difference between the two guesses is most notable for the Porphyrin model − Fe^{2+} system.

Algorithms based on $g_{A,B}$ iteration functions. Comparing the results in Tables 3 and 4, we observe that DIIS algorithms as implemented in GAMESS barely benefit from a better initial guess. Four different choices of A_{tt}, B_{tt} coefficients lead to convergence for Pyridine–Fe^{2+} and Porphyrin model − Fe^{2+} systems (two for Pyridine–Fe^{3+}) for DIIS.

Again, the SOSCF second-order method converges whatever the choice of A_{tt}, B_{tt} coefficients, in less than 100 iterations (thanks to the improved starting guess) except for two specific choices of coefficients (106 iterations needed with Binkley and Faegri coefficients, for Pyridine–Fe^{2+} system).

Algorithms based on the g_{new} iteration function. Both the DIIS and the ODA converge for all the four systems. As in the previous case, the ODA algorithm is very slow to converge to chemical accuracy and ODA followed by g_{new}+DIIS provides satisfactory convergence results.

5 Conclusion and perspectives

In this article, we have provided a geometrical derivation of the ROHF equations in the density matrix and molecular orbital formalisms. A fundamental aspect of that derivation is, for both formalisms, the characterization of the tangent space of the manifold of ROHF states at a critical point of the ROHF energy functional, as well as its orthogonal complement (for the Frobenius inner product). This analysis lead us to introduce a new, parameter-free, iteration function g_{new} (see Eq. (3.8)), as an alternative to Roothaan-like iteration functions $g_{A,B}$ based on the construction of a (non-physical) effective Hamiltonian $H_{A,B}$, where $A = (A_{dd}, A_{ss}, A_{dd})$ and $B = (B_{dd}, B_{ss}, B_{dd})$ collect six real empirical parameters. An important conceptual difference of the proposed new SCF algorithm with respect to previous works is that it is not based on the usual technique of diagonalization of Fock-like Hamiltonians which can lead to numerical instabilities when the *Aufbau* principle is not fulfilled. Thanks to its geometrical formulation, the present algorithm avoids the ambiguity of the orbital energies for which the Koopman theorem does not apply in the case of the ROHF framework.

The numerical results we report seem to indicate that the DIIS algorithm based on the usual g_{AB} framework with the Guest and Saunders ($A_{tt} = B_{tt} = \frac{1}{2}$) and Roothaan ($A_{tt} = \left(-\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right)$, $B_{tt} = \left(\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}\right)$) iteration functions are quite robust and converge in a reasonable number of iterations, even when starting from the core initial guess. However, these observations, made on a small number of test cases (the ones reported in this paper plus a dozen of other challenging cases), do not guarantee that this algorithm will perform well for all systems and basis sets. Remarkably, the DIIS acceleration has to be enabled from the first iteration to guaranty convergence, which does not correspond to the default setting in most quantum chemistry codes, where DIIS is activated only when close enough to a local minimum.

The numerical results reported here based on our new parameter-free iteration function g_{new} are encouraging as the latter converge for all but a single systems tested in this work, which involves different open-shell transition metal ions interacting with aromatic ligands. The algorithms based on the parameter-free iteration function g_{new} may then provide a useful alternative to the $g_{A,B}$ iteration functions for very challenging systems. In particular, the ODA (involving g_{new}) seems to

be extremely robust and efficient in the early iterations, to reach the attraction basin of a local minimizer. Using ODA for the first few iterations, followed by g_{new}+DIIS is a good candidate for a robust black-box SCF routine.

Acknowledgements

The authors would like to thank Michael F. Herbst, Antoine Levitt, Filippo Lipparini and Tommaso Nottoli for fruitful discussions.This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810367).

Supplementary Material for reproductibility

The Supplementary Material contains the atomic coordinates of the benchmark systems studied in the article. The research code used to produce the numerical data is available at https://github.com/LaurentVidal95/ROHFToolkit.

Appendix A: First-order optimality conditions in MO formalism.

As seen in Section 2, the manifold of ROHF states in MO formalism is the quotient manifold

$$\mathcal{M}_{\text{MO}} = \text{St}(N_o;\mathbb{R}^{N_b})/(\mathcal{O}_{N_d}\times\mathcal{O}_{N_s}).$$

In DM formalism, $\mathcal{M}_{\text{DM}}$ is embedded in $\mathcal{V}_{\text{DM}}$, so that the tangent space of $\mathcal{M}_{\text{DM}}$ at a point x can be directly identified with a subspace of $\mathcal{V}_{\text{DM}}$ (see Fig. 2). Unfortunately, this is not the case for the quotient $\mathcal{M}_{\text{MO}}$. Following [2], a way around the problem (valid for general quotient manifolds) is to identify the tangent space $T_{[\![C_o]\!]}\mathcal{M}_{\text{MO}}$ at given equivalence class $[\![C_o]\!]$ with a subspace of $T_{C_o}\text{St}(N_o;\mathbb{R}^{N_b})$, called the horizontal tangent space at C_o to the manifold $\text{St}(N_o;\mathbb{R}^{N_b})$, and denoted $T^{\text{h}}_{C_o}\text{St}(N_o;\mathbb{R}^{N_b})$. We therefore start by computing the expression of the tangent spaces $T_{C_o}\text{St}(N_o;\mathbb{R}^{N_b})$.

Tangent spaces of $\text{St}(N_o;\mathbb{R}^{N_b})$. Let $C_o = (C_d, C_s) \in \text{St}(N_o;\mathbb{R}^{N_b})$. The orthonormality condition ${C_o}^T C_o = I_{N_o}$ translates on C_d and C_s as $C_d^T C_d = I_{N_d}$, $C_s^T C_s = I_{N_s}$ and $C_d^T C_s = 0$. This writes at first order for a perturbation $z = (D_d|D_s) \in \mathbb{R}^{N_b\times N_o}$

$$\begin{aligned} C_d^T D_d + D_d^T C_d = 0 \quad (1)\\ C_s^T D_s + D_s^T C_s = 0 \quad (2)\\ C_d^T D_s + D_d^T C_s = 0 \quad (3). \end{aligned}$$

Let C_v be the orthogonal complement of C_o such that $C = (C_d|C_s|C_v) \in \mathcal{O}_{N_b}$, and let us decompose D_d and D_s in the basis C:

$$\begin{aligned} D_d &= C_d(D_d^d)^T + C_s(D_d^s)^T + C_v(D_d^v)^T\\ D_s &= C_d(D_s^d)^T + C_s(D_s^s)^T + C_v(D_s^v)^T. \end{aligned} \tag{5.1}$$

Then from (1) and (2), there exists $A_d \in \mathbb{R}^{N_d\times N_d}_{\text{skew}}$ and $A_s \in \mathbb{R}^{N_s\times N_s}_{\text{skew}}$ such that $(D_d^d)^T = A_d$ and $(D_s^s)^T = A_s$. Now (3) writes

$$\begin{aligned} C_d^T(C_sA_s + C_d(D_s^d)^T + C_v(D_s^v)^T) + (-A_dC_d + D_d^sC_s^T + D_d^vC_v^T)C_s = D_d^s + (D_s^d)^T = 0\\ \Leftrightarrow D_d^s = -(D_s^d)^T. \end{aligned}$$

We deduce that for all $C_o = (C_d|C_s)$, the tangent space $T_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ is made of all $z = (D_d|D_s) \in \mathbb{R}^{N_b\times N_o}$ such that

$$D_d = C_d A_d + C_s X^T + C_v Y^T \quad \text{and} \quad D_s = -C_d X + C_s A_s + C_v Z^T \tag{5.2}$$

where $X \in \mathbb{R}^{N_d\times N_s}$, $Y \in \mathbb{R}^{N_d\times N_v}$ and $Z \in \mathbb{R}^{N_s\times N_v}$. This also abbreviate as

$$z = C \begin{pmatrix} A_d & -X & -Y \\ X^T & A_s & -Z \\ Y^T & Z^T & 0 \end{pmatrix} \begin{pmatrix} I_{N_o} \\ 0 \end{pmatrix}. \tag{5.3}$$

Horizontal tangent space. Now let $\pi : \mathrm{St}(N_o;\mathbb{R}^{N_b}) \to \mathcal{M}_{\mathrm{MO}}$ be the canonical projection on $\mathcal{M}_{\mathrm{MO}}$

$$\forall\, C_o \in \mathrm{St}(N_o;\mathbb{R}^{N_b}) \quad \pi(C_o) = [\![C_o]\!].$$

We define the vertical tangent space $T^{\mathrm{v}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ at C_o as

$$T^{\mathrm{v}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}) = T_{C_o}\pi^{-1}([\![C_o]\!]).$$

and the horizontal tangent space $T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ as its orthogonal complement for the MO scalar product $\langle C_o|C_o'\rangle = \mathrm{Tr}(C_o^T C_o')$:

$$T_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}) = T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}) \oplus T^{\mathrm{v}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}).$$

Intuitively, $T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ only contains the directions of $T_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ that allow escape the equivalence class $[\![C_o]\!]$, so that one has the important property [2]

$$T_{[\![C_o]\!]}\mathcal{M}_{\mathrm{MO}} \simeq T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}). \tag{5.4}$$

Following the same procedure as for $T_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})$ and $T_x\mathcal{M}_{\mathrm{DM}}$, we can show that

$$T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}) = \left\{(C_s X^T + C_v Y^T| - C_d X + C_v Z^T) \text{ where } X \in \mathbb{R}^{N_d\times N_s},\, Y \in \mathbb{R}^{N_d\times N_v}, Z \in \mathbb{R}^{N_s\times N_v}\right\} \tag{5.5}$$

$$T^{\mathrm{v}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b}) = \left\{(C_d A_d|C_s A_s) \text{ where } A_d \in \mathbb{R}^{N_d\times N_d}_{\mathrm{skew}},\, A_s \in \mathbb{R}^{N_s\times N_s}_{\mathrm{skew}}\right\} \tag{5.6}$$

First order optimality conditions. From (5.4) and (5.5) the first order optimality conditions write in MO formalism as

$$\nabla\mathcal{E}(C_{o*}) \in \mathcal{T}^{\mathrm{h}}_{C_{o*}}\mathrm{St}(N_o;\mathbb{R}^{N_b})^{\perp}. \tag{5.7}$$

A straigthforward computation shows that for all $C_o = (C_d|C_s)$, the ambiant gradient for the standard Frobenius scalar product writes

$$\nabla\mathcal{E}(C_o) = (4F_d C_d|4F_s C_s). \tag{5.8}$$

It now remains to find $\mathcal{T}^{\mathrm{h}}_{C_o}\mathcal{M}^{\perp}_{\mathrm{MO}}$. Once again consider $C_o = (C_d|C_s) \in \mathrm{St}(N_o;\mathbb{R}^{N_b})$ and C_v be such that $C = (C_d|C_s|C_v) \in \mathcal{O}_{N_b}$. For all $W = (W_d|W_s) \in \mathcal{V}_{\mathrm{MO}}$, decomposing W on C as in (5.1) yields

$$\begin{aligned} W \in T^{\mathrm{h}}_{C_o}\mathrm{St}(N_o;\mathbb{R}^{N_b})^{\perp} &\Leftrightarrow \begin{cases} \mathrm{Tr}(X^T(W_d^s - (W_s^d)^T) + Y^T W_d^v + Z^T W_s^v) = 0, \\ \forall X \in \mathbb{R}^{N_d\times N_s},\, Y \in \mathbb{R}^{N_d\times N_v},\, Z \in \mathbb{R}^{N_s\times N_v} \end{cases} \\ &\Leftrightarrow \begin{cases} W_d^v = W_s^v = 0 \\ W_d^s = (W_s^d)^T. \end{cases} \\ &\Leftrightarrow \begin{cases} \exists M_d \in \mathbb{R}^{N_d\times N_d},\, M_s \in \mathbb{R}^{N_s\times N_s}, X \in \mathbb{R}^{N_d\times N_s} \\ \text{ such that } W = (C_d M_d^T + C_s X^T|C_d X + C_s M_s^T). \end{cases} \end{aligned}$$

Using (5.7) and (5.8), there exists $M_d \in \mathbb{R}^{N_d \times N_d}$, $M_s \in \mathbb{R}^{N_s \times N_s}$ and $X \in \mathbb{R}^{N_d \times N_s}$ such that

$$4F_{d*}C_{d*} = C_{d*}M_d^T + C_{s*}X^T \text{ and } 4F_{s*}C_{s*} = C_{d*}X + C_{s*}M_s^T. \tag{5.9}$$

Multiplying both expression by C_d^T or C_s^T we obtain

$$M_d = 4C_{d*}^T F_{d*} C_{d*}, \quad M_s = 4C_{s*}^T F_{s*} C_{s*}, \quad X = 2C_{d*}^T (F_{d*} + F_{s*}) C_{s^*} \tag{5.10}$$

so that the optimality conditions finally write

$$\begin{cases} F_{d*}C_{d*} = C_{d*}\left(C_{d*}^T F_{d*} C_{d*}\right) + \dfrac{1}{2} C_{s*}\left(C_{s*}^T (F_{d*} + F_{s*}) C_{d*}\right) \\ F_{s*}C_{s*} = C_{s*}\left(C_{s*}^T F_{s*} C_{s*}\right) + \dfrac{1}{2} C_{d*}\left(C_{d*}^T (F_{d*} + F_{s*}) C_{s*}\right). \end{cases} \tag{5.11}$$

Appendix B: List of local minima

We provide here the energies at convergence for each system, algorithm, and initial guess. Table 6 corresponds to the energies associated to the results of Table 3 while Table 7 corresponds to the energies associated to the results of Table 4. Finally, Table 8 corresponds to the energies reached by the ODA + g_{new}+DIIS method picture in Table 5

	GAMESS [41]		This work		
	$g_{A,B}$ (3.3) - based methods			g_{new} (3.8) - based methods	
A_{tt}, B_{tt} (see Table 1)	SOSCF	DIIS	DIIS	DIIS	ODA
Pyridine–Fe^{2+}					
Guest and Saunders	-1508.134652	-1508.134652	-1508.014203	-1508.014203	-1508.131670
Roothaan	-1508.016536	-1508.134040	-1508.131670		
Euler	-1508.016536	-1508.016536			
Mc Weeny	-1508.016536				
Binkley	-1508.134652				
Faegri	-1508.016536				
Davidson	-1508.016536				
Canonical I	-1508.134652				
Canonical II					
Pyridine–Fe^{3+}					
Guest and Saunders	-1507.414473	-1507.414091	-1507.411509	-1507.407360	-1507.411509
Roothaan	-1507.414473	-1507.343997	-1507.411889		
Euler	-1507.414473	-1507.414097	-1507.411509		
Mc Weeny	-1507.414473				
Binkley	-1507.414473				
Faegri	-1507.414473				
Davidson	-1507.414473				
Canonical I	-1507.414473				
Canonical II	-1507.414473				
Porphyrin model–Fe^{2+}					
Guest and Saunders	-1940.163309	-1940.513025	-1940.510151		-1940.510191
Roothaan	-1940.163309	-1940.335945	-1940.647646		
Euler	-1940.163309				
Mc Weeny	-1940.163309				
Binkley	-1939.977138				
Faegri	-1939.977138				
Davidson	-1939.977138				
Canonical I	-1940.075387				
Canonical II	-1940.267466				

Table 6: Energies at convergence starting from a core initial guess with 6-31G basis set. The table follows the conventions detailed in the introduction to Section 4. The notation DIIS refers to a DIIS method using f as residual function. The DIIS depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). All energies are expressed in Hartrees.

	GAMESS [41]		This work		
	$g_{A,B}$ (3.3) - based methods			g_{new} (3.8) - based methods	
A_{tt}, B_{tt} (see Table 1)	SOSCF	DIIS	DIIS	DIIS	ODA
Pyridine–Fe^{2+}					
Guest and Saunders	-1508.134652	-1508.013967	-1508.132280	-1508.131670	-1508.131670
Roothaan	-1508.016536	-1507.967145	-1508.131671		
Euler	-1508.134652	-1508.134652	-1508.002054		
Mc Weeny	-1508.134652	-1508.134652	-1508.132280		
Binkley	-1508.134652				
Faegri	-1508.134652				
Davidson	-1508.134652				
Canonical I	-1508.134652				
Canonical II	-1508.134652				
Pyridine–Fe^{3+}					
Guest and Saunders	-1507.414473	-1507.409935	-1507.411510	-1507.411509	-1507.411509
Roothaan	-1507.414473	-1507.409935	-1507.411889		
Euler	-1507.357499				
Mc Weeny	-1507.357499				
Binkley	-1507.414473				
Faegri	-1507.414473				
Davidson	-1507.414473				
Canonical I	-1507.414473				
Canonical II	-1507.357499				
Porphyrin model–Fe^{2+}					
Guest and Saunders	-1940.406548	-1940.513025		-1940.510191	-1940.510191
Roothaan	-1940.406548	-1940.335945	-1940.510151		
Euler	-1940.385615	-1940.513025	-1940.654531		
Mc Weeny	-1940.650207	-1940.513025	-1940.527432		
Binkley	-1940.513025				
Faegri	-1940.513025				
Davidson	-1939.977138				
Canonical I	-1940.513025				
Canonical II	-1940.650207				

Table 7: Energies at convergence starting from an extended Hückel initial guess with 6-31G basis set. The table follows the conventions detailed in the introduction to Section 4. The notation DIIS refers to a DIIS method using f as residual function. The DIIS depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). All energies are expressed in Hartrees.

Table 8: Energies at convergence obtained with a few iterations of ODA, followed by g_{new}+DIIS. The algorithm transitions from ODA to DIIS when the residual norm reaches a tolerance of 10^{-2}. The DIIS depth parameter $m_{\max}$ is fixed to 10 (default value in GAMESS). All energies are expressed in Hartrees.