


Calculating the distance from an electronic wave function to the manifold of Slater determinants through the geometry of Grassmannians

Yuri Alexandre Aoto * and Márcio Fabiano da Silva

Center for Mathematics, Computing and Cognition, Federal University of ABC (UFABC), Santo André, 09210-580 São Paulo, Brazil



(Received 5 August 2020; accepted 15 October 2020; published 6 November 2020)

The set of all electronic states that can be expressed as a single Slater determinant forms a submanifold, isomorphic to the Grassmannian, of the projective Hilbert space of wave functions. We explored this fact by using tools of Riemannian geometry of Grassmannians as described by [Absil *et al.*, *Acta Appl. Math.* **80**, 199 (2004)] to propose an algorithm that converges to a Slater determinant that is a critical point of the overlap function with a correlated wave function. This algorithm can be applied to quantify the entanglement or correlation of a wave function. We show that this algorithm is equivalent to the Newton method using the standard parametrization of Slater determinants by orbital rotations, but it can be more efficiently implemented because the orbital basis used to express the correlated wave function is kept fixed throughout the iterations. We present the equations of this method for a general configuration interaction wave function and for a wave function with up to double excitations over a reference determinant. Applications of this algorithm to selected electronic systems are also presented and discussed.

DOI: [10.1103/PhysRevA.102.052803](https://doi.org/10.1103/PhysRevA.102.052803)

I. INTRODUCTION

Electron correlation is at the heart of electronic structure theory, and its intimate relation to quantum entanglement, as viewed by quantum information theory, attracts the attention of researchers from both fields [1–3]. From the point of view of atomic and molecular physics, the effect of electron correlation on the electronic energy is the most important feature to be considered, although its consequences on properties are also relevant in several applications. From the side of quantum information theory, one is often interested in quantifying the entanglement of a wave function intrinsically [3–5], irrespective of any observable, in particular the energy. Many ways to measure entanglement have been proposed [4], for instance, by the distance between the quantum state and the set of states with no entanglement (uncorrelated states). This definition has a geometric nature, and calls for the geometry of the sets of quantum states.

In wave-function methods of electronic structure, the correlation due to the fermionic character of the electrons is always taken into account by using antisymmetrized wave functions. Slater determinants, which represent mean-field states, are the simplest of such wave functions. The true ground-state wave function, on the other hand, presents extra electron correlation apart of that associated with the Pauli principle, and if one is concerned with measuring this extra correlation or entanglement in the wave function, its distance to the set of Slater determinants is an expected approach [2]. Using the set of configuration state functions, that are spin eigenfunctions, is a similar possibility, although more involved.

There are multiple ways to define a metric in the space of electronic states. It cannot be a metric on the Hilbert space of wave functions, but on its projective space instead, as it must reflect the fact that the normalization and phase of the wave functions do not alter their associated physical states. Some examples of metrics used in quantum mechanics are

$$D_{\text{FS}}(\Psi, \Psi') = \arccos|\langle\Psi|\Psi'\rangle|, \quad (1)$$

$$D_{\text{ACFC}}(\Psi, \Psi') = \sqrt{1 - |\langle\Psi|\Psi'\rangle|}, \quad (2)$$

$$D_{\text{BRLCM}}(\Psi, \Psi') = 1 - |\langle\Psi|\Psi'\rangle|^2, \quad (3)$$

where the wave functions Ψ and Ψ' are assumed to be normalized to unit. Equation (1) is the Fubini-Study metric, introduced in quantum mechanics by Bures [6,7], and it can be interpreted as the angle between both state vectors. The second equation has been studied by D'Amico *et al.* [8], together with a related metric in the space of densities (their original metric uses a different normalization condition). The last one was recently used by Benavides-Riveros *et al.* [2], who provided an upper bound to $D_{\text{BRLCM}}(\Phi_{\text{HF}}, \Psi_0)$ based on the correlation energy (for systems with nondegenerate ground state):

$$\begin{aligned} 1 - |\langle\Phi_{\text{HF}}|\Psi_0\rangle|^2 &= D_{\text{BRLCM}}(\Phi_{\text{HF}}, \Psi_0) \\ &\leq \frac{|E_{\text{corr}}|}{E_{\text{gap}}} = \frac{E_{\text{HF}} - E_0}{E_1 - E_0}, \end{aligned} \quad (4)$$

where Φ_{HF} and Ψ_0 are the Hartree-Fock and the exact ground-state wave functions, respectively, and E_0 , E_1 , and E_{HF} are the ground state, first excited state, and Hartree-Fock energies. Furthermore, quantification of static and dynamic correlation incorporated by a wave function based on this metric has also been proposed [9].

*yuri.aoto@ufabc.edu.br

For all the above metrics, the distance between two states described by wave functions Ψ and Ψ' is related to the absolute value of their overlap $|\langle\Psi|\Psi'\rangle|$ by Eqs. (1) to (3). Therefore, calculating the distance from a given correlated electronic wave function to the set of Slater determinants is equivalent to finding the Slater determinant that maximizes the overlap with such wave function (constrained to normalized wave functions). This is not a trivial task [3], and analytical expressions are known only for the small case of two-particle systems [10] and for some specific cases [11]. In [2], Benavides-Riveros *et al.* opted to measure the correlation of the ground-state wave function by its distance to the Hartree-Fock wave function [in Eq. (4)], that is not, in general, the minimizer of the distance among all possible Slater determinants. Since no analytical procedure is known for the general case, finding the Slater determinant that minimizes the distance to a correlated wave function requires a numerical optimization on the space of Slater determinants. This optimization problem has been studied by Zhang and Kollar [10] and by Zhang and Mauser [11], by both analytical and numerical procedures. They have also provided an algorithm that converges monotonically, but slowly, to the Slater determinant with the largest overlap with a correlated wave function. Their algorithm is discussed below in Sec. III C 3.

To fully appreciate the phenomenon of correlation in electronic structure, one has to consider how the manifold of Slater determinants is embedded in the set of all electronic wave functions. This manifold is the *Grassmannian* [12,13]. Although Slater determinants are of paramount importance in electronic structure theory and Grassmannians are of high importance to geometry [14,15], their connection is rarely observed when looking at Slater determinants. The first work to establish this connection dates back to 1980, by Rowe, Ryman, and Rosensteel [16]. Afterwards, Cassam-Chenaï (in 1994) [17], Panin (in 2007) [18], and Chiumiento and Melgaard (in 2012) [19] have also studied the geometry of the Grassmannian in such context. Very recently, Polack and coworkers have used the Grassmannian to formulate a procedure to obtain an initial guess for self-consistent field calculations [20]. However, systematic applications of the properties of this “fundamental family of compact complex manifolds” [14] to the electronic structure theory is still missing.

The objective of this article is to explore how the geometry of the Grassmannian can be used to perform the optimization of the Slater determinant with the largest overlap with an arbitrary wave function, and thus ultimately measuring its correlation. We will show that if we explicitly consider the geometry of the Grassmannian, a more efficient algorithm can be obtained. Moreover, this algorithm is a Newton method that uses a set of *nonindependent parameters* to describe the Slater determinants (namely, the coefficient matrices on a fixed orbital basis), contrary to the usual assumption that a set of independent parameters is necessary to carry out orbital optimizations with the Newton method.

In Sec. II we review the relation between the Grassmannian and the field of electronic structure, along with the Plücker embedding and the description of many-electron wave functions by the exterior algebra [21,22]. Up to our knowledge, researchers in molecular physics and theoretical

chemistry usually have no familiarity with these concepts. For the detailed treatment of electronic wave functions within the framework of exterior algebra, the reader is referred to the works of Muddin [23,24], and Cassam-Chenaï [17]. See also [25] for applications of this approach in the context of quantum computation. Section II is ended with the mathematical formulation of the problem we will be concerned with. Section III describes the algorithms we propose and compares them from the theoretical point of view. In Sec. IV some numerical examples are discussed. After the concluding remarks, we present in the Appendix the complete expressions for these algorithms over symmetry-adapted spatial orbitals.

II. GRASSMANNIAN AND THE PLÜCKER EMBEDDING

For a molecular system of n electrons, a finite-dimensional approximation for the space of wave functions can be obtained by first fixing a finite-dimensional vector space of one-electron wave functions (the *spin-orbital space*):

$$\mathcal{W} = \text{span}\{\phi_p\}_{p=1}^M. \quad (5)$$

This M -dimensional vector space is usually defined by choosing a basis set for an electronic structure calculation. From the spin-orbital space one can construct n -electron wave functions as linear combinations of n -electron Slater determinants made by elements of \mathcal{W} . These n -electron wave functions form the required vector space, that is, the n th exterior power of \mathcal{W} , denoted by $\bigwedge^n \mathcal{W}$ [21]. From the point of view of the exterior algebra, an n -electron Slater determinant is the *exterior product*, or *wedge product* \wedge , of n elements of \mathcal{W} . For instance,

$$\begin{aligned} |\Phi\rangle &= \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \dots & \varphi_n(1) \\ \varphi_1(2) & \varphi_2(2) & \dots & \varphi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(n) & \varphi_2(n) & \dots & \varphi_n(n) \end{vmatrix} \\ &= \varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n, \end{aligned} \quad (6)$$

where $\{\varphi_i\}_{i=1}^n$ is a linearly independent set, but otherwise arbitrary, of elements of \mathcal{W} . Therefore, the vector space $\bigwedge^n \mathcal{W}$ consists of all possible linear combinations of such elements, and it is identified with the n -electron sector of the Fock space (whereas the complete Fock space is identified with the *exterior algebra* of \mathcal{W} , $\bigwedge \mathcal{W} = \bigoplus_{n=0}^M \bigwedge^n \mathcal{W}$). In this article, *ket* notation will be used for n -electron wave functions, whereas small Greek letters are used for one-electron wave functions (orbitals). Aside from subscripts, Slater determinants will be generally denoted by $|\Phi\rangle$ and arbitrary n -electron wave functions by $|\Psi\rangle$.

Elements such as $\varphi_1 \wedge \dots \wedge \varphi_n \in \bigwedge^n \mathcal{W}$, that can be written as the exterior product of elements of \mathcal{W} , are said to be *decomposable* (the nomenclatures *simple* and *free* are also used by some authors, the latter especially in the context of quantum entanglement [3]). Slater determinants are thus the decomposable elements of $\bigwedge^n \mathcal{W}$. Furthermore, given a basis for \mathcal{W} , such as in Eq. (5), the set of all decomposable elements made by n elements of this basis (with no repetition and taken in order, e.g., by ascending indices) forms a basis for $\bigwedge^n \mathcal{W}$.

An arbitrary element of $\bigwedge^n \mathcal{W}$ can be constructed as

$$\begin{aligned} |\Psi\rangle &= \sum_{I, |I|=n} C_I \phi_{I_1} \wedge \cdots \wedge \phi_{I_n} \\ &= \sum_{I, |I|=n} C_I a_{I_1}^\dagger \cdots a_{I_n}^\dagger | \rangle \\ &= \sum_{I, |I|=n} C_I |\Phi_I\rangle, \end{aligned} \quad (7)$$

where the summation runs over all multi-indices sets I with length n . For convenience, the indication $|I| = n$ will be often dropped, as the length of the multi-indices set is always the number of electrons, that will be clear by the context. Equation (7) is clearly interpreted as a configuration interaction (CI) expansion. We also make the connection with the formalism of second quantization, with a^\dagger being creation operators and $| \rangle$ the vacuum state. Note that $|\Psi\rangle$ might be decomposable or not, depending whether there is a basis for \mathcal{W} , say $\{\phi'_p\}_{p=1}^M$, such that

$$|\Psi\rangle = \phi'_1 \wedge \cdots \wedge \phi'_n. \quad (8)$$

In general, given a wave function in the form of Eq. (7), it is not evident if it is decomposable. We want to find a characterization of the set of all wave functions in $\bigwedge^n \mathcal{W}$ that are decomposable.

Consider the following Slater determinant:

$$|\Phi\rangle = \varphi_1 \wedge \cdots \wedge \varphi_n, \quad (9)$$

with each φ_i being an element of \mathcal{W} . If the set $\{\varphi_1, \dots, \varphi_n\}$ is linearly dependent, the Slater determinant vanishes, namely, it is the zero element of the vector space $\bigwedge^n \mathcal{W}$. Assuming that this set is linearly independent, it spans a n -dimensional vector subspace of \mathcal{W} :

$$\text{span}\{\varphi_1, \dots, \varphi_n\} \subset \mathcal{W}, \quad (10)$$

$$\dim(\text{span}\{\varphi_1, \dots, \varphi_n\}) = n. \quad (11)$$

Obviously, this vector space admits infinitely many other bases, obtained from $\{\varphi_1, \dots, \varphi_n\}$ by nonsingular linear transformations:

$$\varphi'_i = \sum_{p=1}^M \varphi_p U_i^p, \quad (12)$$

$$\text{span}\{\varphi_1, \dots, \varphi_n\} = \text{span}\{\varphi'_1, \dots, \varphi'_n\}. \quad (13)$$

Furthermore, and of central importance for the present argument, the decomposable element made by φ'_i differs to $|\Phi\rangle$ [Eq. (9)] by normalization or phase only:

$$\varphi'_1 \wedge \cdots \wedge \varphi'_n = \lambda \varphi_1 \wedge \cdots \wedge \varphi_n, \quad \lambda \neq 0. \quad (14)$$

Thus, there is a one-to-one map between the physical states that can be represented by a Slater determinant and the set of n -dimensional subspaces of \mathcal{W} . The set of all n -dimensional vector subspaces of a given vector space \mathcal{W} is the *Grassmannian*, or the *Grassmann manifold*. It will be represented by $\text{Gr}(n, \mathcal{W})$.

Recall that the normalization and phase of a wave function is not relevant for the description of the physical state it represents. Thus, the space of states for the n -electron system

is actually the *projective space* of $\bigwedge^n \mathcal{W}$, denoted by $\mathbb{P} \bigwedge^n \mathcal{W}$. This space is the set of the equivalence classes in $\bigwedge^n \mathcal{W}$ obtained by the relation $|\Psi\rangle \sim \lambda |\Psi\rangle$, where λ is a nonzero scalar. The equivalence class of $|\Psi\rangle$, indicated by $[|\Psi\rangle]$, is composed by the wave functions that differ from $|\Psi\rangle$ by a normalization or phase factor.

From the above discussion, an element of $\text{Gr}(n, \mathcal{W})$ (a vector subspace of \mathcal{W}) is associated to the equivalence class of a Slater determinant (an element of $\mathbb{P} \bigwedge^n \mathcal{W}$). This is a bijection with the set of all Slater determinants (except for a scalar factor), that forms a submanifold in the space of wave functions that is a copy of the Grassmannian $\text{Gr}(n, \mathcal{W})$ inside $\mathbb{P} \bigwedge^n \mathcal{W}$:

$$\text{Slater determinants} \leftrightarrow n\text{-electron wave functions}, \quad (15)$$

$$\text{Gr}(n, \mathcal{W}) \leftrightarrow \mathbb{P} \bigwedge^n \mathcal{W}, \quad (16)$$

$$[|\Phi\rangle] \mapsto [\varphi_1 \wedge \cdots \wedge \varphi_n]. \quad (17)$$

The application given in Eq. (17) is known as the *Plücker embedding* [14,26]. The image of $\text{Gr}(n, \mathcal{W})$ in $\mathbb{P} \bigwedge^n \mathcal{W}$ satisfies a set of quadratic equations in $\mathbb{P} \bigwedge^n \mathcal{W}$, the *Plücker relations*, of high importance in the field of algebraic and projective geometries. Thus, an element $|\Psi\rangle \in \bigwedge^n \mathcal{W}$ is decomposable if and only if its coefficients on a basis made by decomposable elements [as in Eq. (7)] satisfy the Plücker relations. A particular case of these relations will be shown in Sec. II B.

The strong connection between Slater determinants and the Grassmannian suggests that the properties of the latter can be used to work with the former, in particular for their optimization. This will be explored in the remaining of this article. Thus, we will make no distinction between a wave function, that is a point in $\bigwedge^n \mathcal{W}$, and its equivalence class in $\mathbb{P} \bigwedge^n \mathcal{W}$; we will often say that a wave function is at the Grassmannian, or belongs to the Grassmannian, when it can be represented by a Slater determinant; we will also interchange the nomenclatures, and make no distinction between decomposable elements of $\bigwedge^n \mathcal{W}$, Slater determinants, and the vector subspace of \mathcal{W} spanned by its orbitals:

$$\begin{aligned} [|\Phi\rangle] &= [\varphi_1 \wedge \cdots \wedge \varphi_n] \\ &= \text{span}\{\varphi_1, \dots, \varphi_n\} \in \text{Gr}(n, \mathcal{W}) \subset \mathbb{P} \bigwedge^n \mathcal{W}. \end{aligned} \quad (18)$$

A. Representation of the Grassmannian

The most obvious way to represent a Slater determinant is by a $M \times n$ matrix of rank n , denoted by U , having the coefficients of a basis of the Slater determinant in a fixed basis of \mathcal{W} [e.g., of Eq. (5)]:

$$|\Phi\rangle = \phi'_1 \wedge \cdots \wedge \phi'_n \Leftrightarrow U \in \mathbb{R}^{M \times n}, \quad (19)$$

$$\phi'_i = \sum_{p=1}^M \phi_p U_i^p. \quad (20)$$

Thus, column i has the M coefficients of orbital ϕ'_i on this basis of \mathcal{W} . Because U is of rank n , its columns are linearly independent and span an n -dimensional vector space, associated to the n -electron Slater determinant. We write $[|\Phi\rangle] = \text{span}(U)$. This matrix is not unique [see Eq. (13)].

Thus, to run over all the Grassmannian when looking for an optimal Slater determinant of any sort, the entries of U cannot be freely varied because a change on the entries of U might provoke no change on the corresponding Slater determinant, or lead to a matrix U with linearly dependent columns (that does not span an n -dimensional vector space and does not represent an n -electron Slater determinant).

Slater determinants can also be parametrized starting from a pivot Slater determinant, say,

$$|\Phi_0\rangle = \phi_1 \wedge \cdots \wedge \phi_n, \quad (21)$$

by [16,27,28]

$$|\Phi\rangle = e^{\hat{K}}|\Phi_0\rangle, \quad (22)$$

where

$$\hat{K} = \sum_{i=1}^n \sum_{a=n+1}^M K_i^a (a_a^\dagger a_i - a_i^\dagger a_a). \quad (23)$$

A transformation matrix from this basis to a basis of $|\Phi\rangle$ is given by

$$U_{\text{full}} = \exp \left\{ \begin{pmatrix} \mathbf{0}_{n \times n} & -\mathbf{K}^T \\ \mathbf{K} & \mathbf{0}_{(M-n) \times (M-n)} \end{pmatrix} \right\}. \quad (24)$$

Note that the matrix U in Eq. (19) corresponds to the first n columns of U_{full} .

This second parametrization is originated from the works of Thouless [27], and has been studied by several authors in the context of electronic structure theory [16,29–31]. It offers a set of $n(M-n)$ independent parameters K_i^a (not considering possible symmetry constraints), that is exactly the dimension of the Grassmannian [12]. This parametrization is largely used to carry out variations on the orbitals in self-consistent field (SCF) methods [28,32–37], and one often refers to occupied-virtual (and other “type X–type Y”) orbital rotations in orbital optimizations. It is the *de facto* parametrization used in modern SCF calculations. However, in Sec. III B we will show that the first type of representation described above, Eq. (19), can also be used for a Newton optimization of Slater determinants, contrary to what is often assumed [28,35,36].

B. Example: Hydrogen molecule

We will illustrate the concepts discussed so far for the $M_S = 0$ states of the hydrogen molecule H_2 , described by a minimal basis set ($1s$ alpha and beta orbitals centered in each atom), with real coefficients and orbitals. This case allows a visualization of the Grassmannian as embedded in the projective space of the two-electron wave functions, depicted in Fig. 1. For convenience, we will use symmetry-adapted orbitals:

$$\phi_+ = N_+(1s_A + 1s_B), \quad (25)$$

$$\phi_- = N_-(1s_A - 1s_B), \quad (26)$$

with $N_{\pm} = \frac{1}{\sqrt{2(1 \pm \langle 1s_A | 1s_B \rangle)}}$, but the present discussion does not depend by any means on this particular basis, and only in Sec. IV A it will become apparent the reason for this choice.

$$\begin{array}{ll} \phi_+ \wedge \bar{\phi}_+ & \bullet \\ \phi_+ \wedge \bar{\phi}_- & \blacksquare \\ \phi_- \wedge \bar{\phi}_+ & \blacktriangle \\ \phi_- \wedge \bar{\phi}_- & \blacklozenge \end{array}$$

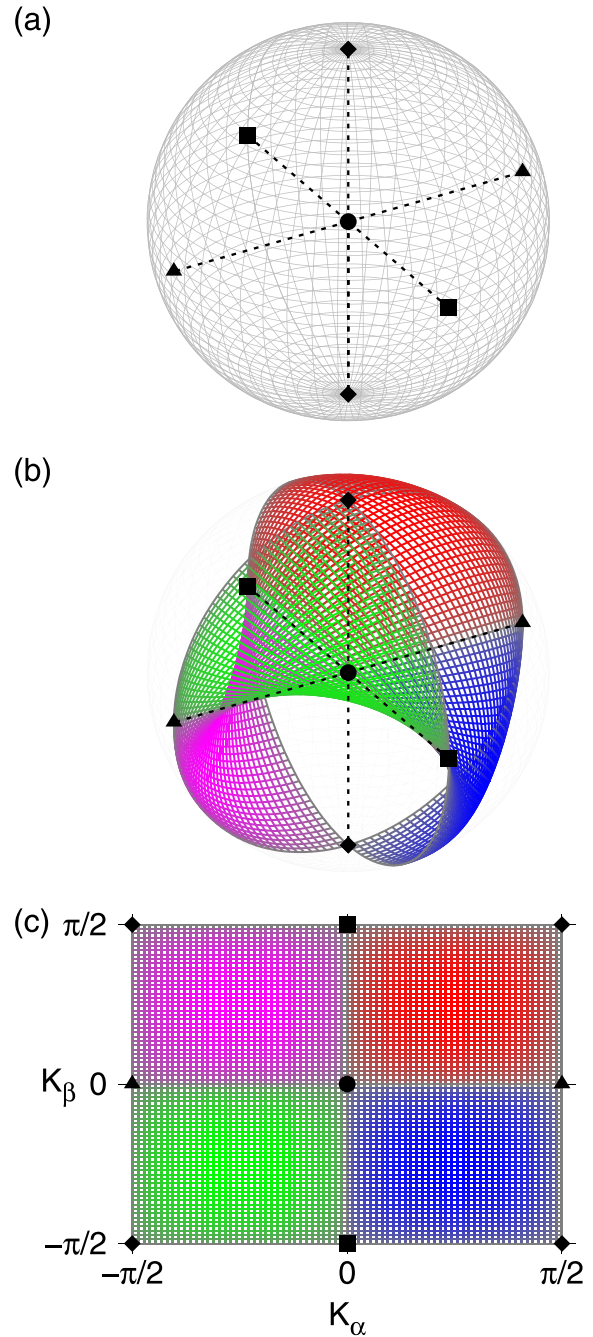


FIG. 1. (a) A representation of $\mathbb{P}(\wedge^2 \mathcal{W}_{H_2})_{M_S=0}$, along with the four Slater determinants given in Eq. (29). (b) The $M_S = 0$ subset of the Grassmannian $\text{Gr}(2, \mathcal{W}_{H_2})$ embedded in $\mathbb{P}(\wedge^2 \mathcal{W}_{H_2})_{M_S=0}$. (c) The space of the parameters K_α and K_β , that determines the surface at (b) by Eq. (38).

The orbital and two-electron wave function spaces are (beta-spin orbitals are indicated by overlines)

$$\mathcal{W}_{H_2} = \text{span}\{\phi_+, \phi_-, \bar{\phi}_+, \bar{\phi}_-\}, \quad (27)$$

$$\begin{aligned} \bigwedge^2 \mathcal{W}_{H_2} = \text{span}\{ & \phi_+ \wedge \overline{\phi_+}, \phi_+ \wedge \overline{\phi_-}, \\ & \phi_- \wedge \overline{\phi_+}, \phi_- \wedge \overline{\phi_-}, \\ & \phi_+ \wedge \phi_-, \overline{\phi_+} \wedge \overline{\phi_-}\}, \end{aligned} \quad (28)$$

$$\begin{aligned} \left(\bigwedge^2 \mathcal{W}_{H_2} \right)_{M_S=0} = \text{span}\{ & \phi_+ \wedge \overline{\phi_+}, \phi_+ \wedge \overline{\phi_-}, \\ & \phi_- \wedge \overline{\phi_+}, \phi_- \wedge \overline{\phi_-}\}. \end{aligned} \quad (29)$$

As $(\bigwedge^2 \mathcal{W}_{H_2})_{M_S=0}$ is four dimensional, its projective space is three dimensional. An arbitrary element of $(\bigwedge^2 \mathcal{W}_{H_2})_{M_S=0}$ is

$$\begin{aligned} |\Psi\rangle = & C_{13} \phi_+ \wedge \overline{\phi_+} + C_{14} \phi_+ \wedge \overline{\phi_-} + C_{23} \phi_- \wedge \overline{\phi_+} \\ & + C_{24} \phi_- \wedge \overline{\phi_-}. \end{aligned} \quad (30)$$

A visual representation of this space can be obtained in the following way: The elements of $\mathbb{P}(\bigwedge^2 \mathcal{W}_{H_2})_{M_S=0}$ can be viewed as the ‘‘rays’’ (straight lines that pass through the origin) in a four-dimensional space. Each of these lines cross the unit sphere $S^3 \subset \mathbb{R}^4$ in two antipodal points. They represent the two normalized wave functions in $(\bigwedge^2 \mathcal{W}_{H_2})_{M_S=0}$ associated to the same physical state, but differing by sign. To fix one representative element, we choose the one with non-negative coordinate for (say) $\phi_+ \wedge \overline{\phi_+}$, that is, the ‘‘upper hemisphere’’ of S^3 with respect to the direction of $\phi_+ \wedge \overline{\phi_+}$. Finally, we project these points of $S^3 \subset \mathbb{R}^4$ into the \mathbb{R}^3 space defined by $C_{13} = 0$, obtaining the unit ball in \mathbb{R}^3 , so that each of its points represents an element of $\mathbb{P}(\bigwedge^2 \mathcal{W}_{H_2})_{M_S=0}$. This representation is depicted in Fig. 1(a). It can be interpreted as what an observer in \mathbb{R}^4 would see, when looking at S^3 from the top. At the center of the visualization ‘‘plane’’ there is $[\phi_+ \wedge \overline{\phi_+}]$ (as this observer is looking exactly from its direction), and the other elements of the basis in Eq. (30) are represented (twice) in the boundary of the ball.

The Plücker relation that characterizes how the Grassmannian $\text{Gr}(2, \mathcal{W}_{H_2})$ is embedded in $\mathbb{P}(\bigwedge^2 \mathcal{W}_{H_2})$ is (there are much more equations, with more terms, for larger cases) [14]

$$C_{12}C_{34} - C_{13}C_{24} + C_{14}C_{23} = 0, \quad (31)$$

where C_{12} and C_{34} are the coefficients of Slater determinants with $M_S \neq 0$, and are zero in the present case. Furthermore, we are choosing normalized wave functions with $C_{13} \geq 0$, and thus the relation becomes

$$C_{24} \sqrt{1 - (C_{24}^2 + C_{14}^2 + C_{23}^2)} - C_{14}C_{23} = 0. \quad (32)$$

The Grassmannian is represented by the set of points in the space of variables $\{C_{24}, C_{14}, C_{23}\}$ that satisfies this relation, and the corresponding wave functions can be written as single Slater determinants. Figure 1(b) illustrates how this manifold is embedded in $\mathbb{P}(\bigwedge^2 \mathcal{W}_{H_2})$, with the representation described above.

Observe that the straight lines seen in Figs. 1(a) and 1(b), joining $\phi_+ \wedge \overline{\phi_+}$ with each one of the other Slater determinants in the basis of Eq. (30), represent wave functions that are linear combinations of $\phi_+ \wedge \overline{\phi_+}$ with only one other element of Eq. (30). For instance, the vertical line represents wave functions of the form

$$|\Psi\rangle = C_{13} \phi_+ \wedge \overline{\phi_+} + C_{24} \phi_- \wedge \overline{\phi_-}. \quad (33)$$

In the ‘‘north hemisphere’’, C_{13} and C_{24} have the same sign, whereas in the ‘‘south hemisphere’’ they have opposite signs. Except for the cases when one of C_{13} or C_{24} is zero, the points of this line are not at the Grassmannian [see Fig. 1(b)], what means that the wave function above cannot in general be taken as a single Slater determinant, a well-known fact in electronic structure theory. On the other hand, the horizontal straight lines, connecting $\phi_+ \wedge \overline{\phi_+}$ and $\phi_+ \wedge \overline{\phi_-}$ or $\phi_+ \wedge \overline{\phi_-}$, are at the Grassmannian. The corresponding wave functions are

$$\begin{aligned} |\Phi_1\rangle = & C_{13} \phi_+ \wedge \overline{\phi_+} + C_{23} \phi_- \wedge \overline{\phi_+}, \\ |\Phi_2\rangle = & C_{13} \phi_+ \wedge \overline{\phi_+} + C_{14} \phi_+ \wedge \overline{\phi_-}, \end{aligned} \quad (34)$$

that are linear combinations of $\phi_+ \wedge \overline{\phi_+}$ with single excitations from ϕ_+ to ϕ_- (at alpha or beta spin, respectively). As it turns out to be, linear combinations with singly excited determinants can always be represented as a single Slater determinant:

$$\begin{aligned} |\Phi_1\rangle = & (C_{13} \phi_+ + C_{23} \phi_-) \wedge \overline{\phi_+}, \\ |\Phi_2\rangle = & \phi_+ \wedge (C_{13} \overline{\phi_+} + C_{14} \overline{\phi_-}). \end{aligned} \quad (35)$$

The fact that $|\Psi\rangle$ [in Eq. (33)] does not belong to the Grassmannian in general, whereas $|\Phi_1\rangle$ and $|\Phi_2\rangle$ do, is clearly seen in the Plücker relations (31) and (32): the coefficients of $|\Psi\rangle$ do not satisfy Eq. (32), but the coefficients of $|\Phi_1\rangle$ and $|\Phi_2\rangle$ do.

Equation (35) suggests that every point of the $M_S = 0$ subset of the Grassmannian $\text{Gr}(2, \mathcal{W}_{H_2})$ can be obtained as

$$|\Phi\rangle = (a \phi_+ + b \phi_-) \wedge (c \overline{\phi_+} + d \overline{\phi_-}), \quad (36)$$

to be represented by the matrix

$$U = \begin{pmatrix} a & 0 \\ b & 0 \\ 0 & c \\ 0 & d \end{pmatrix}. \quad (37)$$

The division into blocks comes from the fact that we are concerned with a single value for M_S , what naturally excludes mixing among alpha and beta orbitals. On the other hand, the orbital $a \phi_+ + b \phi_-$ can also be interpreted as a rotation of ϕ_+ toward ϕ_- by an angle of $K_\alpha = \arctan(\frac{b}{a})$. Analogously, one defines $K_\beta = \arctan(\frac{d}{c})$. The corresponding normalized Slater determinant, made of normalized orbitals, becomes

$$\begin{aligned} |\Phi\rangle = & [\cos(K_\alpha) \phi_+ + \sin(K_\alpha) \phi_-] \\ & \wedge [\cos(K_\beta) \overline{\phi_+} + \sin(K_\beta) \overline{\phi_-}] \\ = & e^{\hat{K}} (\phi_+ \wedge \overline{\phi_+}), \end{aligned} \quad (38)$$

where \hat{K} is defined in Eq. (23). Equations (37) and (38) are the two parametrizations of the Grassmannian, based on the coefficient matrices and on orbital rotations, described in Sec. II A. Finally, the $M_S = 0$ subset of the Grassmannian is complete if the space of the parameters K_α and K_β is $[-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$, as shown in Fig. 1(c).

The concepts of this section are illustrated in an animation in the Supplemental Material [38]. An analogous representation of the $M_S = 0$ subset of the Grassmannian has been obtained by Cassam-Chenai [17], although he represents the

projective space in a more pictorial fashion. The reader is strongly referred to his work for a deeper discussion of the concepts presented in this section.

C. Overlap with a correlated wave function

Consider now an arbitrary wave function $|\Psi_{\text{ext}}\rangle \in \bigwedge^n \mathcal{W}$, that does not necessarily belong to the Grassmannian (subscript “ext” stands for “external to the Grassmannian”). For instance, this could be the exact or some approximate wave function for the ground state of the system. The objective of this work is to devise and study algorithms to find the Slater determinant that maximizes the overlap to this wave function. This makes sense only for the equivalence classes of $|\Psi_{\text{ext}}\rangle$ and of the Slater determinants, their corresponding elements in $\mathbb{P}\bigwedge^n \mathcal{W}$. Thus, our problem consists in optimizing the following function, defined at the Grassmannian:

$$f(|\Phi\rangle) = \frac{|\langle\Phi|\Psi_{\text{ext}}\rangle|}{\sqrt{\langle\Phi|\Phi\rangle\langle\Psi_{\text{ext}}|\Psi_{\text{ext}}\rangle}}. \quad (39)$$

III. ALGORITHMS

In this section we will discuss algorithms for the optimization of the function defined in Eq. (39), using the Newton method. To simplify the analysis and the discussion of the equations, they will be presented on a spin-orbital basis, with no inclusion of spatial symmetry. Complete equations using spatial orbitals and considering spatial symmetry from Abelian point groups are given in the Appendix.

A. Algorithm 1: By independent parameters in $|\Phi\rangle = e^{\mathbf{K}}|\Phi_0\rangle$

With the parametrization given by Eq. (22) one can directly apply the Newton method. Let

$$\begin{aligned} f_{\square} : \mathbb{R}^{n(M-n)} &\rightarrow \mathbb{R} \\ \mathbf{K} &\mapsto f(e^{\mathbf{K}}|\Phi_0\rangle) \end{aligned} \quad (40)$$

be the function that represents the overlap function f [Eq. (39)], but defined over the space of orbital rotation parameters. An improved Slater determinant is obtained from $|\Phi_0\rangle$ with the parameters K_i^a (collected in \mathbf{K}) that solve the equation

$$\mathfrak{H}\mathbf{K} = -\mathfrak{J}, \quad (41)$$

where \mathfrak{J} and \mathfrak{H} , the Jacobian and Hessian of f_{\square} , are made by the first and second derivatives of f_{\square} . Calculation of f_{\square} , \mathfrak{J} , and \mathfrak{H} is straightforward when $\mathbf{K} = 0$, and $|\Psi_{\text{ext}}\rangle$ is given as a normalized linear combination of excitations on top of $|\Phi_0\rangle$:

$$|\Psi_{\text{ext}}\rangle = C_0|\Phi_0\rangle + \sum_{i,a} C_i^a|\Phi_i^a\rangle + \frac{1}{4} \sum_{i,j,a,b} C_{ij}^{ab}|\Phi_{ij}^{ab}\rangle + \dots, \quad (42)$$

with $C_{ij}^{ab} = C_{ji}^{ba} = -C_{ij}^{ba} = -C_{ji}^{ab}$. In such conditions,

$$f_{\square}(\mathbf{K} = 0) = C_0, \quad (43)$$

$$\mathfrak{J}_i^a = \frac{\partial f_{\square}(\mathbf{K} = 0)}{\partial K_i^a} = (-1)^{i+n} C_i^a, \quad (44)$$

$$\mathfrak{H}_{ij}^{ab} = \frac{\partial^2 f_{\square}(\mathbf{K} = 0)}{\partial K_i^a \partial K_j^b} = \begin{cases} -C_0, & a = b, i = j \\ -(-1)^{i+j} C_{ij}^{ab}, & \text{otherwise.} \end{cases} \quad (45)$$

Note that the Jacobian is constructed from the coefficients of single excitations, whereas the Hessian is formed by the coefficients of the reference (diagonal elements) and of the double excitations (all with respect to $|\Phi_0\rangle$). However, these expressions are valid only if $\mathbf{K} = 0$, otherwise, higher-rank excitations also contribute, in a nontrivial way. Hence, if this algorithm is used to optimize f , the basis of \mathcal{W} used to expand $|\Psi_{\text{ext}}\rangle$ has to be changed in every iteration, to obtain the coefficients as excitations with respect to the new Slater determinant $|\Phi\rangle = e^{\mathbf{K}}|\Phi_0\rangle$.

Changing the orbital basis used to expand $|\Psi_{\text{ext}}\rangle$ is very disadvantageous. First of all, because this is a time-consuming step. If carried out in a straightforward way, it is accomplished by the formula

$$\begin{aligned} |\Psi_{\text{ext}}\rangle &= \sum_I C_I |\Phi_I\rangle = \sum_I C_I \phi_{I_1} \wedge \dots \wedge \phi_{I_n} \\ &= \sum_I C'_I |\Phi'_I\rangle = \sum_I C'_I \phi'_{I_1} \wedge \dots \wedge \phi'_{I_n}, \end{aligned} \quad (46)$$

$$C'_I = \sum_J C_J \det((U_{\text{full}}^J)_I), \quad (47)$$

where U_{full} is the transformation matrix from the basis $\{\phi'_p\}$ to the basis $\{\phi_p\}$, obtained from \mathbf{K} by Eq. (24). The matrices $(U_{\text{full}}^J)_I$ are the minors of the matrix U_{full} , with the entries in the rows and columns given by the multi-indices I and J . This is the bottleneck step, and its computational cost is discussed in Sec. III C 2. However, the main disadvantage is that the representation of the external wave function is changed in every iteration. This implies that, if $|\Psi_{\text{ext}}\rangle$ is an approximate wave function based on some kind of rank truncation, say, a configuration interaction with single and double excitations (CISD) wave function over the reference $|\Phi_0\rangle$, it will not contain only single and double excitations over the new Slater determinant $|\Phi\rangle = e^{\mathbf{K}}|\Phi_0\rangle$. Therefore, the initial *structure* of this wave function is lost, and a full configuration interaction (FCI) like wave function has always to be used as external wave function. Throughout this text, any possible rank truncation scheme used to construct $|\Psi_{\text{ext}}\rangle$ will be generally denoted as the *structure* of $|\Psi_{\text{ext}}\rangle$.

B. Algorithm 2: Using the coefficients matrix U

The difficulties in the algorithm above arise from the standard parametrization by orbital rotations [Eq. (22)]. It is often assumed that a set of independent parameters is necessary to perform a Newton optimization of orbitals, as otherwise the Hessian matrix is singular or near singular close to the optimal orbitals [28]. This would exclude the possibility of using directly the matrix U , defined in Eq. (19), for such kind of optimization. However, this is exactly what the optimization procedure on the Grassmannian of Absil and coworkers do [39].

Using the $M \times n$ matrices of full rank to represent the elements of the Grassmannian, Absil and coworkers have

studied the Riemannian geometry of the Grassmannian. These matrices form the noncompact Stiefel manifold $ST(n, M)$, and all computations are carried out on it, whereas tools of differential geometry are used to go back and forth from the Stiefel manifold to the Grassmann manifold. With this technique, Absil and coworkers presented several formulas for geometric concepts on Grassmannians, such as canonical metric and geodesics. More important for this work, a Newton method specific for the Grassmann manifold was also proposed. It works in the following way [39]: Given $f : Gr(n, \mathcal{W}) \rightarrow \mathbb{R}$ a function defined on the Grassmannian with real values sufficiently smooth, let

$$f_{\diamond} : ST(n, M) \rightarrow \mathbb{R} \\ U \mapsto f(\text{span}(U)) \quad (48)$$

be the corresponding function on the Stiefel manifold, that is, the function defined over $M \times n$ matrices that, when calculated on any representative matrix U of $|\Phi_0\rangle$, returns $f(|\Phi_0\rangle)$. The procedure is carried out by computations over the matrix U , that belongs to the Stiefel manifold:

(i) One first solves the following equation for the unknown $\eta_{\diamond U} \in H_U = \{U_{\perp} K : K \in \mathbb{R}^{(M-n) \times n}\}$:

$$\Pi_{U_{\perp}} D(\Pi_{\perp} \text{grad} f_{\diamond}(\cdot))(U)[\eta_{\diamond U}] = -\Pi_{U_{\perp}} \text{grad} f_{\diamond}(U). \quad (49)$$

(ii) Update $|\Phi_0\rangle \rightarrow |\Phi\rangle$ by moving along the geodesic on the Grassmannian in the direction of $\eta_{\diamond U}$, by computing a singular value decomposition (SVD) of $\eta_{\diamond U} = \mathcal{U} \Sigma \mathcal{V}^T$ and calculating

$$[|\Phi\rangle] = \text{span}(U \mathcal{V} \cos \Sigma + \mathcal{U} \sin \Sigma). \quad (50)$$

In these equations, U_{\perp} is any full-rank $M \times (M - n)$ matrix such that $U^T U_{\perp} = 0$; the gradient of f_{\diamond} at U is the $M \times n$ matrix whose entries are given by $(\text{grad} f_{\diamond}(U))_q^p = \frac{\partial f_{\diamond}(U)}{\partial U_q^p}(U)$; $DF(x)[y] = \frac{d}{dt} F(x + ty)|_{t=0}$ is the directional derivative of F at x in the direction of y ; $\Pi_{W_{\perp}} = I - W(W^T W)^{-1} W^T$ is the projection onto the orthogonal complement of the matrix W . The center dot in Eq. (49) denotes the point where the function has to be evaluated (and thus it stands for $U + t\eta_{\diamond U}$ when calculating the directional derivative); and the matrices \mathcal{U} , Σ , and \mathcal{V} are $M \times n$ orthonormal, $n \times n$ diagonal, and $n \times n$ orthonormal, respectively.

There is a clear analogy between Eq. (49) and the standard Newton method, Eq. (41): the left-hand side of Eq. (49) is related to the second derivatives of f_{\diamond} , calculated in the direction of $\eta_{\diamond U}$, whereas in the right-hand side are the first derivatives. However, the projectors onto the orthogonal complement of U and the requirement that the unknown $\eta_{\diamond U}$ belongs to H_U (for every element η of H_U , $U^T \eta = 0$) guarantee that variations in U that do not change $|\Phi_0\rangle$ are canceled out [39]. Thus, the usage of a redundant set of parameters does not pose a problem here.

We have adapted this procedure for the overlap function f , defined in Eq. (39). Assuming U orthogonal, $U^T U = \mathbb{1}$, this function can be calculated as

$$f(|\Phi\rangle) = f_{\diamond}(U) = \sum_I C_I F_I, \quad (51)$$

and Eq. (49) becomes the following system of linear equations:

$$\sum_{r=1}^M \sum_{s=1}^n \mathcal{H}_{qs}^{pr} (\eta_{\diamond U})_s^r = -\mathcal{J}_q^p, \quad (52)$$

where

$$\mathcal{H}_{qs}^{pr} = \sum_{\bar{p}=1}^M (\Pi_{U_{\perp}})_{\bar{p}}^p \sum_I C_I (\tilde{\mathbf{H}}_I)_{qs}^{\bar{p}r}, \quad (53)$$

$$\mathcal{J}_q^p = \sum_{\bar{p}=1}^M (\Pi_{U_{\perp}})_{\bar{p}}^p \sum_I C_I (\mathbf{G}_I)_q^{\bar{p}}, \quad (54)$$

$$(F_I) = \det(U|_I), \quad (55)$$

$$(\mathbf{G}_I)_q^p = \det((U \stackrel{q}{\leftarrow} e_p)|_I), \quad (56)$$

$$(\mathbf{H}_I)_{qs}^{pr} = \det((U \stackrel{q}{\leftarrow} e_p \stackrel{s}{\leftarrow} e_r)|_I), \quad (57)$$

$$(\tilde{\mathbf{H}}_I)_{qs}^{pr} = \begin{cases} (\mathbf{H}_I)_{qs}^{pr} & \text{if } s \neq q, \\ -F_I \delta_{pr} & \text{otherwise.} \end{cases} \quad (58)$$

The notation $A|_I$ indicates the $n \times n$ submatrix of A whose rows are in the multi-index I , $A \stackrel{q}{\leftarrow} b$ represents the matrix A with the q th column replaced by b , and e_p is the p th element of the canonical basis of \mathbb{R}^M . The condition $\eta_{\diamond U} \in H_U$ can be imposed by extending the linear system with the n^2 equations $U^T \eta_{\diamond U} = 0$, and solving it with a least-square subroutine.

To see how the quantities \mathbf{G}_I and \mathbf{H}_I appear, note that

$$\begin{aligned} \frac{\partial F_I}{\partial U_q^p}(U) &= \frac{\partial \det(U|_I)}{\partial U_q^p}(U) \\ &= \text{tr} \left(\text{adj}(U|_I) \frac{\partial (U|_I)}{\partial U_q^p}(U) \right) \\ &= \text{tr} \begin{pmatrix} \det(U \stackrel{1}{\leftarrow} \delta_{1q} e_p)|_I & \dots & \det(U \stackrel{1}{\leftarrow} \delta_{nq} e_p)|_I \\ \vdots & \ddots & \vdots \\ \det(U \stackrel{n}{\leftarrow} \delta_{1q} e_p)|_I & \dots & \det(U \stackrel{n}{\leftarrow} \delta_{nq} e_p)|_I \end{pmatrix} \\ &= (\mathbf{G}_I)_q^p, \end{aligned} \quad (59)$$

where $\text{adj}(W)$ represents the classical adjoint, or adjugate, of the matrix W . The quantities \mathbf{H}_I appear similarly from the directional derivative of the gradient in the left-hand side of Eq. (49). Thus, the matrices \mathbf{G}_I and \mathbf{H}_I are associated to the first and second derivatives of f_{\diamond} . The factor $-F_I$ in the diagonal entries of $\tilde{\mathbf{H}}_I$ is a contribution from the projector Π_{\perp} to the directional derivative [see Eq. (49)].

One has to calculate several $n \times n$ determinants in this algorithm, as it is necessary for the basis transformation of the wave function $|\Psi_{\text{ext}}\rangle$ [see Eq. (46)]. However, in the case of the present algorithm the number of such determinants is smaller (see Sec. III C 2) and the external wave function is expressed always in the same basis, with the CI coefficients that appear in the equations being the same at all iterations.

The case of a CISD wave function and the usage of symmetry-adapted spatial orbitals

We apply the equations outlined above for the case where the external wave function is a CISD (configuration

interaction with single and double excitations) wave function $|\Psi_{\text{CISD}}\rangle$. This exemplifies the advantages of this algorithm: first, the summations over I in Eq. (52) are always over the same single and double excitations, with the same CI coefficients throughout the iterations. This does not happen for the algorithm based on orbital rotations, described in Sec. III A. In that case, the wave function $|\Psi_{\text{CISD}}\rangle$ must be given as a linear combination of excitations on top of the Slater determinant of each iteration, that obviously changes, and it is not necessarily the reference used to construct the CISD wave function.

A second advantage is that an efficient implementation is possible, that benefits from the structure of the wave function, in particular by exploiting symmetry-adapted spatial orbitals. In such case the matrix U is divided into blocks and so are the matrices used to calculate F_I , \mathbf{G}_I , and \mathbf{H}_I . Because the determinant of a block-diagonal matrix is the product of the determinant of each block, these quantities are decomposed in similar quantities for each spin and irreducible representation. For example, if all indices $pqrs$ belong to the same spin and irreducible representation, say alpha orbitals and irreducible representation $\Gamma = 1$, then

$$(\mathbf{H}_I)_{qs}^{pr} = (\mathbf{H}_{I_\Gamma}^{\alpha})_{qs}^{pr} \prod_{\Gamma \neq 1} F_{I_\Gamma}^{\alpha} \prod_{\Gamma} F_{I_\Gamma}^{\beta}, \quad (60)$$

whereas if indices pq belong to alpha orbitals, but rs belong to beta orbitals (and same irreducible representation $\Gamma = 1$):

$$(\mathbf{H}_I)_{qs}^{pr} = (\mathbf{G}_{I_\Gamma}^{\alpha})_q^p (\mathbf{G}_{I_\Gamma}^{\beta})_s^r \prod_{\Gamma \neq 1} F_{I_\Gamma}^{\alpha} F_{I_\Gamma}^{\beta}. \quad (61)$$

The complete equations for spin-restricted CISD wave functions are given and discussed in the Appendix.

C. Theoretical comparison of the algorithms

1. Procedures are equivalent

For every initial Slater determinant $|\Phi_0\rangle$, both algorithms produce the same Newton step $|\Phi_0\rangle \rightarrow |\Phi\rangle$. In fact, first note that Algorithm 2, based on the coefficients matrices U , takes into account the intrinsic geometry of the Grassmannian as studied by Absil and coworkers [39]. Thus, it is independent of the chosen orbital basis, either for the complete one-particle space \mathcal{W} or for the vector space $[|\Phi_0\rangle]$. Therefore, once an orbital basis for $|\Phi_0\rangle$ has been chosen,

$$|\Phi_0\rangle = \phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_n, \quad (62)$$

one can always consider a basis for \mathcal{W} that contains it, and extend it to the virtual space of $|\Phi_0\rangle$:

$$\mathcal{W} = [\phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_n \wedge \cdots \wedge \phi_M]. \quad (63)$$

In this condition, the matrix that represents $|\Phi_0\rangle$ is trivial:

$$U = \begin{pmatrix} \mathbf{1}_{n \times n} \\ \mathbf{0}_{(M-n) \times n} \end{pmatrix}. \quad (64)$$

Now, let $|\Psi_{\text{ext}}\rangle$ be expanded in the basis $\{\phi_p\}_{p=1}^M$, such that both algorithms can be applied. This is not required for Algorithm 2 but we consider it here for the sake of the present argument. A straightforward application of Eq. (52) with U given by Eq. (64) shows that the Jacobian and the Hessian in

TABLE I. Number of distinct $n \times n$ determinants used in both algorithms, without considering any spin or spatial symmetry restriction. The number of electrons is represented by n , and the total number of spin orbitals by K . The total number of possible Slater determinants is N_{full} , whereas N represents the number of Slater determinants with nonvanishing contribution to the wave function in question. N_{full} is assumed to grow as $\frac{1}{n} \left(\frac{K}{n}\right)^K$ [28].

Alg. 1		$N_{\text{full}}^2 = \binom{K}{n}^2 = \left(\frac{K!}{n!(K-n)!}\right)^2 \sim \frac{1}{n^2} \left(\frac{K}{n}\right)^{2K}$
Alg. 2	general	$N[1 + nK + (nK)^2] \sim Nn^2K^2$
	FCI	$\left(\frac{K!}{n!(K-n)!}\right)[1 + nK + (nK)^2] \sim \left(\frac{K}{n}\right)^K nK^2$
	CISD	$[1 + n(K-n) + \frac{n(n-1)(K-n)(K-n-1)}{4}][1 + nK + (nK)^2] \sim n^4K^4$

the Algorithm 1, based on orbital rotations, appear as submatrices at the right- and left-hand side matrices:

$$\mathcal{J} = \begin{pmatrix} \mathbf{0}_{n \times n} \\ \mathfrak{J} \end{pmatrix}, \quad (65)$$

$$\mathcal{H}_q^p = \begin{pmatrix} \mathbf{a}_{n \times n} \\ \mathfrak{H}_q^p \end{pmatrix}. \quad (66)$$

Submatrix \mathbf{a} , although nonzero, does not affect the solution of Eq. (52) since all columns of $\eta_{\diamond U} \in H_U$ must be orthogonal to the columns of U [$\eta_{\diamond U} \in H_U$, see comment before Eq. (49)]:

$$\eta_{\diamond U} = \begin{pmatrix} \mathbf{0}_{n \times n} \\ \mathbf{K} \end{pmatrix} = \mathcal{U}\Sigma\mathcal{V}^T, \quad (67)$$

where we recall that a singular value decomposition will be applied to $\eta_{\diamond U}$ in Algorithm 2. Submatrix \mathbf{K} in Eq. (67) is the same that solves Eq. (41) of Algorithm 1. Finally, by expanding the functions exponential, cosine, and sine for matrices, one shows that the updated U matrix obtained by Eq. (50) is the same as the first n columns of Eq. (24), except by multiplication to the orthogonal matrix \mathcal{V}^T (from the singular value decomposition), that does not change the corresponding Slater determinant:

$$[|\Phi\rangle] = [e^{\hat{K}}|\Phi_0\rangle] = \text{span}(U\mathcal{V} \cos \Sigma + \mathcal{U} \sin \Sigma). \quad (68)$$

Hence, both procedures are equivalent.

2. Computational cost of the algorithms

The algorithms presented above scale differently with the system size. They both rely on the calculation of a large number of $n \times n$ determinants, with n being the number of electrons [compare Eq. (47) with Eqs. (55)–(57)]. However, the number of such determinants is quite different in each case, and they are compared in Table I. Algorithm 1 scales exponentially, irrespective of the kind of external wave function, since a full orbital transformation has to be performed. On the other hand, in Algorithm 2, the basis used to expand the external wave function is preserved along the iterations, and thus the number of Slater determinants used for its expansion remains unchanged (N). This might be much smaller than the total number of possible Slater determinants (N_{full}), as for a CISD external wave function. However, even for a full configuration interaction (FCI) external wave function the number of

determinant calculations that need to be performed increases slower than in Algorithm 1, although still exponentially. In actual computations, the external wave function is usually eigenfunction of S_z , and often considers molecular spatial symmetry, reducing the number of determinant calculations. Relations among F_I , \mathbf{G}_I , and \mathbf{H}_I of similar multi-indices can also be used to avoid calculating several of the determinants, speeding up computations (see the Appendix A 1 b). Furthermore, the transformation of the wave function in Algorithm 1 can be performed much efficiently by the procedure proposed by Malmqvist [40]. However, this analysis shows that a faster procedure can be obtained with Algorithm 2, in particular because it can exploit the structure of the external wave function.

3. Comparison with the algorithm of Zhang and Kollar

A possible pitfall of the Newton method as applied here is that it might converge to a saddle point of the overlap function, as it strongly depends on the initial Slater determinant (this is further explored in Sec. IV A). In our test applications, discussed in Sec. IV, this has not been a problem since we are concerned with ground-state wave functions, for which the restricted Hartree-Fock wave function is a perfectly fine starting guess. Furthermore, for these cases the present algorithm converges quite fast, typically in three iterations. On the other hand, the algorithm proposed by Zhang and Kollar [10] is more robust, optimizing the overlap function by working with one orbital at a time: in each iteration, all orbitals of the Slater determinant are fixed except for, say, ϕ_i ; this orbital is updated such that the overlap function is maximized under the restriction that ϕ_i is orthogonal to all other orbitals of the Slater determinant. In the next iteration the procedure is repeated with ϕ_{i+1} (now with ϕ_i fixed), or back to ϕ_1 , cyclically, until convergence is obtained. This procedure is guaranteed to converge to a maximum, although not necessarily to a global maximum [10]. However, it takes a large number of iterations to converge, especially after reaching the plateau where just small updates are made after each iteration (see Fig. 1 of [10]). Thus, the present algorithms, based on the Newton method, and the one from Zhang and Kollar, are complementary and could be used in conjunction: the more robust, but slower, algorithm of Zhang and Kollar can be used to reach the region of the Grassmannian close to the optimum point (where iterations lead to small variations of the overlap function), and thereafter the Newton method, as presented here, used for a fast convergence toward the critical point.

IV. SOME EXAMPLES

We have coded pilot implementations of the algorithms discussed in the previous section by using PYTHON. Tensor contractions to generate the elements of Eq. (52) are straightforwardly implemented with NUMPY [41,42]. A hand-crafted C/Fortran code can obviously speed up computations, but the present implementations suffice for our initial purposes. For Algorithm 2, based on coefficient matrices, two implementations have been coded: one for a general $|\Psi_{\text{ext}}\rangle$ and one specific for $|\Psi_{\text{ext}}\rangle$ being a CISD wave function. The accuracy of the implementations has been checked by the following means:

(1) The Jacobian and Hessian in the Algorithm 1 have also been calculated by finite differences and compared to the analytical versions [Eqs. (44) and (45)].

(2) We performed sanity checks on the Algorithm 2 to ensure that the solution of Eq. (52), $\eta_{\diamond U}$, really satisfies the original Eq. (49), of Absil and coworkers, with derivatives calculated by finite differences.

(3) Both implementations of the Algorithm 2 lead to the same matrices of Eq. (52).

(4) All implementations give the same iterations when the same $|\Psi_{\text{CISD}}\rangle$ is used, as required by the conclusion of Sec. III C 1.

In this section we will describe some example calculations carried out with these implementations.

A. Hydrogen molecule in a minimal basis

We start by searching the Slater determinant with largest overlap to the exact ground-state wave function (the FCI wave function $|\Psi_0\rangle$) for the hydrogen molecule described by a minimal basis set, as discussed in Sec. II B. The STO-3G basis-set representation [43] has been used. Analytic expressions to measure entanglement and correlation can be derived for this case [44]. It is still of profound physical importance since the distance between $|\Psi_0\rangle$ and the Grassmannian reaches its maximum in the dissociation limit, with noninteracting electrons, as recently discussed by Ding and Schilling [3]. Here, we will focus on the behavior of the overlap function.

For this small case there is a simple expression for the function to be optimized, being clear where its maximum is: The exact wave function has the form of Eq. (33) and, from Eq. (38), one obtains (considering normalized wave functions)

$$\langle \Phi(K_\alpha, K_\beta) | \Psi_0 \rangle = C_0 \cos(K_\alpha) \cos(K_\beta) + \sqrt{1 - C_0^2} \sin(K_\alpha) \sin(K_\beta), \quad (69)$$

where C_0 is the coefficient of $\phi_+ \wedge \overline{\phi_+}$ (the ‘‘reference determinant’’), and $\sqrt{1 - C_0^2}$ is the coefficient of $\phi_- \wedge \overline{\phi_-}$ (the ‘‘doubly excited determinant’’). The absolute value of this overlap assumes its maximum at $\phi_+ \wedge \overline{\phi_+}$ ($K_\alpha = K_\beta = 0$), if $|C_0| > \frac{1}{\sqrt{2}}$. This is the case for every internuclear distance R , in particular close to the equilibrium distance, where the weight of the reference determinant is much larger than of the excited determinant (see Fig. 2). However, note that $\phi_- \wedge \overline{\phi_-}$ ($|K_\sigma| = \pi/2$) is another critical point of this function, and the Newton method might converge to it, depending on the Slater determinant used to start the optimization. In general, the overlap function might have several critical points over the Grassmannian, and the optimization procedure might not converge to a maximum. In practice, if $|\Psi_0\rangle$ is the exact or an approximate wave function for the ground state, an obvious starting point is the Hartree-Fock Slater determinant.

The behavior of the overlap function depends on the distance between $|\Psi_0\rangle$ and the Grassmannian. If $|\Psi_0\rangle$ is close to the Grassmannian, the maximum at $(K_\alpha, K_\beta) = (0, 0)$ is very clear. When $|\Psi_0\rangle$ becomes far from the Grassmannian, moving away from $\phi_+ \wedge \overline{\phi_+}$, it gets closer to opposite regions of the Grassmannian, in particular to $\phi_- \wedge \overline{\phi_-}$, whose contribution to $|\Psi_0\rangle$ increases. In the limit $R \rightarrow \infty$, $|C_0| = \frac{1}{\sqrt{2}}$ and

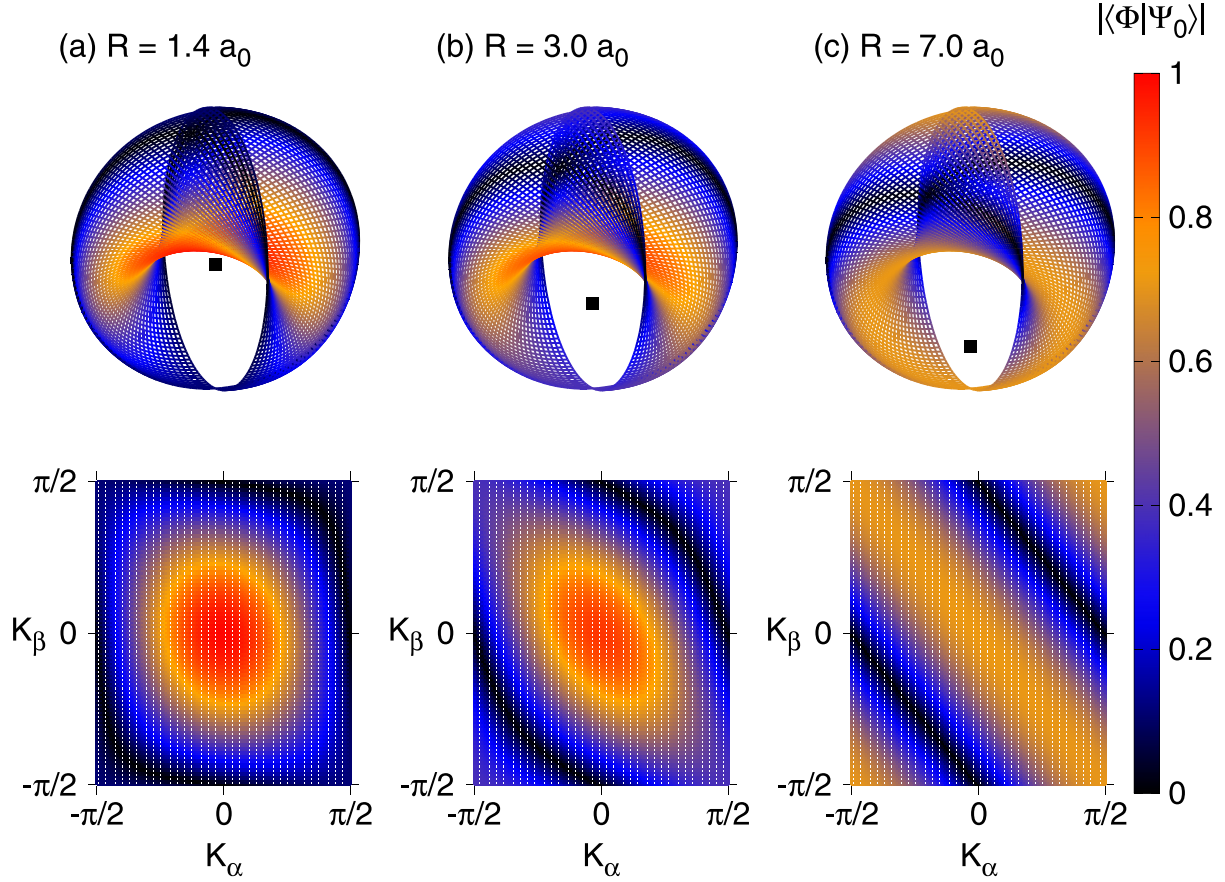


FIG. 2. The function $f_{\square}(K_{\alpha}, K_{\beta}) = |\langle \Phi(K_{\alpha}, K_{\beta}) | \Psi_0 \rangle|$, for the hydrogen molecule, for the internuclear distances of (a) $1.4a_0$, (b) $3.0a_0$, and (c) $7.0a_0$. The upper images show the three-dimensional representation as discussed in Sec. II B, and the bottom images show the space of the parameters K_{α} and K_{β} . The exact wave function $|\Psi_0\rangle$ is represented by a square in the upper images.

$|\Psi_0\rangle$ becomes equally distant to every Slater determinant of the form

$$\begin{aligned}
 |\Phi\rangle &= [\cos(K)\phi_+ + \sin(K)\phi_-] \\
 &\wedge [\cos(-K)\bar{\phi}_+ + \sin(-K)\bar{\phi}_-] \\
 &= (a\phi_+ + \sqrt{1-a^2}\phi_-) \\
 &\wedge (a\bar{\phi}_+ - \sqrt{1-a^2}\bar{\phi}_-). \quad (70)
 \end{aligned}$$

The maximum overlap is reached not at a single point of the Grassmannian, but in a submanifold of it. This submanifold, seen in Fig. 2(c), is the stripe crossing diagonally the parameters space. [Note that, although $|\Psi_0\rangle$ appears to be closer to $\phi_- \wedge \bar{\phi}_-$ in Fig. 2(c), this is just because this representation does not preserve the metric of the projective space.]

B. Selected systems

For H_2 in a minimal basis set ($1s$ orbital centered in each atom), the point in the Grassmannian with maximum overlap with the exact ground-state wave function is just the restricted Hartree-Fock Slater determinant $|\Phi_{\text{HF}}\rangle = \phi_+ \wedge \bar{\phi}_+$ (see Sec. IV A). This is not the case in general. In this section we present the value of $|\langle \Phi_{\text{minD}} | \Psi_{\text{ext}} \rangle|$ for selected systems, where $|\Phi_{\text{minD}}\rangle$ is the optimized Slater determinant with largest overlap with $|\Psi_{\text{ext}}\rangle$ [and thus with minimum distance to $|\Psi_{\text{ext}}\rangle$

according to the metrics in Eqs. (1)–(3)]. As for $|\Psi_{\text{ext}}\rangle$, we use the configuration interaction with single and double excitations (CISD) wave function, calculated with the MOLPRO package [45]. We consider the following basis sets: STO-3G [43], 6-31G [46], and cc-pVnZ [47] ($n \in \{\text{D}, \text{T}, \text{Q}\}$). The optimization procedure is started with $|\Phi_{\text{HF}}\rangle$, the restricted Hartree-Fock wave function. The frozen core approximation has been used in all examples; this implies that core orbitals of $|\Phi_{\text{minD}}\rangle$ are the same of $|\Phi_{\text{HF}}\rangle$ since the elements of the Jacobian associated to these orbitals are zero [see Eq. (44)]. Thus, the optimization of $|\Phi_{\text{minD}}\rangle$ can be made over a Grassmannian of smaller order, by considering only orbitals correlated in $|\Psi_{\text{CISD}}\rangle$.

Aside from applying the present algorithms, the objective of this section is to compare $|\Phi_{\text{minD}}\rangle$ and $|\Phi_{\text{HF}}\rangle$ or, equivalently, $|\langle \Phi_{\text{minD}} | \Psi_{\text{CISD}} \rangle|$ and $|\langle \Phi_{\text{HF}} | \Psi_{\text{CISD}} \rangle|$, the latter being directly available after a CISD calculation. Note that

$$\begin{aligned}
 \langle \Phi_{\text{HF}} | \Psi_{\text{CISD}} \rangle &= \langle \Phi_{\text{HF}} | \hat{P}_{\text{minD}} + \hat{Q}_{\text{minD}} | \Psi_{\text{CISD}} \rangle \\
 &= \langle \Phi_{\text{HF}} | \Phi_{\text{minD}} \rangle \langle \Phi_{\text{minD}} | \Psi_{\text{CISD}} \rangle \\
 &\quad + \langle \Phi_{\text{HF}} | \hat{Q}_{\text{minD}} | \Psi_{\text{CISD}} \rangle \\
 &\approx \langle \Phi_{\text{HF}} | \Phi_{\text{minD}} \rangle \langle \Phi_{\text{minD}} | \Psi_{\text{CISD}} \rangle, \quad (71)
 \end{aligned}$$

where $\hat{P}_{\text{minD}} = |\Phi_{\text{minD}}\rangle \langle \Phi_{\text{minD}}|$ is the projector onto $|\Phi_{\text{minD}}\rangle$ and $\hat{Q}_{\text{minD}} = 1 - \hat{P}_{\text{minD}}$. The approximation in Eq. (71) holds

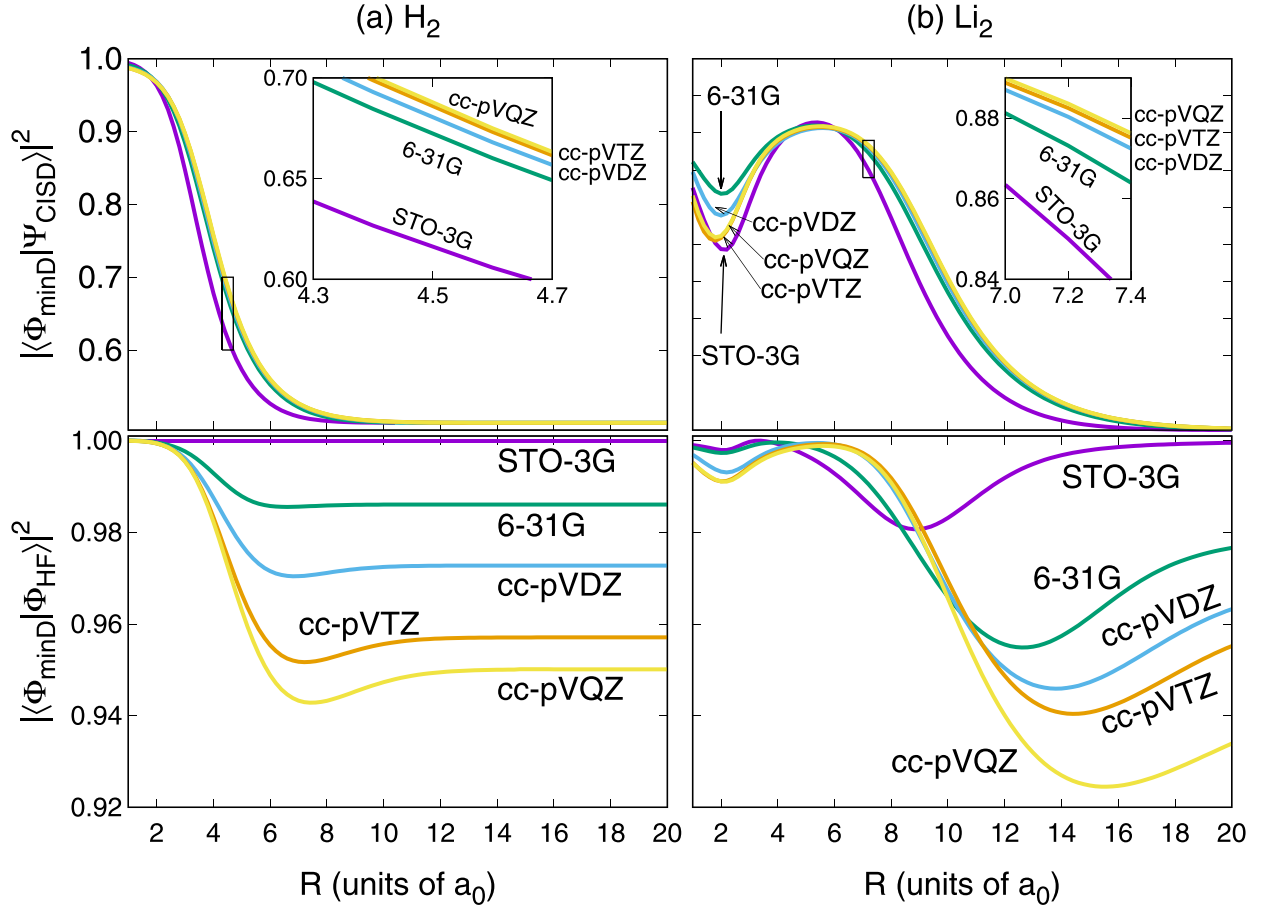


FIG. 3. The square of the absolute value for the overlap between $|\Phi_{\min D}\rangle$, and $|\Psi_{\text{CISD}}\rangle$ (top graphs) or $|\Phi_{\text{HF}}\rangle$ (bottom graphs). Inset graphs show the curves in a closer range at the region indicated by the rectangle.

whenever $|\Phi_{\min D}\rangle$ is not too far from $|\Phi_{\text{HF}}\rangle$, hence the term with $Q_{\min D}$ can be neglected. Thus,

$$\frac{\langle \Phi_{\text{HF}} | \Psi_{\text{CISD}} \rangle}{\langle \Phi_{\min D} | \Psi_{\text{CISD}} \rangle} \approx \langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle. \quad (72)$$

Recall that $|\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle|$ is related to the distance (measured in $\mathbb{P} \wedge^n \mathcal{W}$) between these two Slater determinants, and either side of Eq. (72) can be used to estimate the importance of using the optimized $|\Phi_{\min D}\rangle$ (instead of plain $|\Phi_{\text{HF}}\rangle$) to measure the correlation. In the calculations we carried out, Eq. (72) holds.

The first application is for H_2 in larger basis sets, being $|\Psi_{\text{CISD}}\rangle$ the exact wave function. In Fig. 3(a) we show the values of $|\langle \Phi_{\min D} | \Psi_{\text{CISD}} \rangle|^2$ and $|\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle|^2$ as function of the internuclear distance. The qualitative behavior of $|\langle \Phi_{\min D} | \Psi_0 \rangle|^2$ is the same at all the basis sets, namely, it rapidly decreases during the dissociation, reaching the value of circa 0.5. The reasoning has been discussed in Sec. IV A. However, for larger basis sets, $|\Phi_{\min D}\rangle$ deviates from $|\Phi_{\text{HF}}\rangle$, in particular at larger interatomic distances. For the cc-pVQZ basis set at $R \approx 7.0a_0$, $|\langle \Phi_{\min D} | \Psi_0 \rangle|^2$ is 94% of the value of $|\langle \Phi_{\text{HF}} | \Psi_0 \rangle|^2$.

For the Li_2 molecule, we observe a more complex variation of $|\langle \Phi_{\min D} | \Psi_0 \rangle|^2$ along the dissociation, presenting a maximum at $R \approx 5.5a_0$, slightly after the equilibrium distance (at

$R = 5a_0$). This maximum indicates that the correlation at this point is minimum, as already discussed by Benavides-Riveros and coworkers [9]. Although the qualitative behavior of $|\langle \Phi_{\min D} | \Psi_0 \rangle|^2$ does not change with basis set, $|\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle|^2$ is strongly dependent on the basis set. The deviation of $|\langle \Phi_{\min D} | \Psi_0 \rangle|^2$ from $|\langle \Phi_{\text{HF}} | \Psi_0 \rangle|^2$ increases for larger basis sets, particularly in regions where correlation is also large. Note that, under the frozen core approximation, the CISD wave function is also exact for Li_2 .

Table II shows results from the optimization of $|\Phi_{\min D}\rangle$, with respect to the CISD wave function, for some other molecules. As it happens for the previous examples, $|\Phi_{\min D}\rangle$ is quite close to $|\Phi_{\text{HF}}\rangle$, especially for systems with small correlation, such as water in the equilibrium geometry. For the stretched water molecule correlation effects increase and $|\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle|^2$ decreases, but it is still over 0.98. Observe that $\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle$ is rather less sensitive to the increase of basis set than $\langle \Phi_{\min D} | \Psi_{\text{CISD}} \rangle$. For the ozone molecule, and for the three transition metal diatomic molecules, both $\langle \Phi_{\min D} | \Psi_{\text{CISD}} \rangle$ and $\langle \Phi_{\min D} | \Phi_{\text{HF}} \rangle$ are quite insensitive with basis set.

V. CONCLUSIONS

In this article we described procedures to optimize the critical points of the overlap to an arbitrary wave function over the set of Slater determinants. This can be used to measure the

TABLE II. Results for the optimization of the Slater determinant with the largest overlap with the CISD wave function ($|\Phi_{\text{minD}}\rangle$), for selected systems and basis sets.

Molecule	Geometry	Basis set	$ \langle\Phi_{\text{minD}} \Psi_{\text{CISD}}\rangle ^2 \times 100$	$ \langle\Phi_{\text{minD}} \Phi_{\text{HF}}\rangle ^2 \times 100$
H ₂ O	$R_{\text{OH}} = 0.9633 \text{ \AA}$ $a_{\text{HOH}} = 102.57^\circ$	cc-pVDZ	95.063	99.961
		cc-pVTZ	94.504	99.954
		cc-pVQZ	94.391	99.945
H ₂ O	$R_{\text{OH}} = 2.5 \text{ \AA}$ $a_{\text{HOH}} = 102.57^\circ$	cc-pVDZ	63.356	98.533
		cc-pVTZ	70.812	98.481
		cc-pVQZ	72.786	98.518
O ₃	$R_{\text{OO}} = 1.2728 \text{ \AA}$ $a_{\text{OOO}} = 116.75^\circ$	cc-pVDZ	87.310	99.405
		cc-pVTZ	87.181	99.539
		cc-pVQZ	87.215	99.572
ScH	$R_{\text{ScH}} = 1.7754 a_0$	cc-pVDZ	92.059	99.785
		cc-pVTZ	92.361	99.769
		cc-pVQZ	92.472	99.769
CuH	$R_{\text{CuH}} = 1.4626 a_0$	cc-pVDZ	93.451	99.722
		cc-pVTZ	93.544	99.761
		cc-pVQZ	93.481	99.761
ZnO	$R_{\text{ZnO}} = 1.7047 a_0$	cc-pVDZ	92.016	99.593
		cc-pVTZ	91.916	99.698
		cc-pVQZ	91.827	99.723

distance between a correlated wave function and the set of Slater determinants, that is a measure of correlation incorporated in the wave function. Obtaining such distance is important to understand the relation between electronic correlation and entanglement [2,3], and to analyze the interplay between static and dynamic correlation [9]. The optimization procedures described here can be used for relatively large systems, using exact as well as approximate wave functions. An efficient version specific for a configuration interaction with single and double excitations (CISD) wave function is presented. We developed these procedures by acknowledging that the set of Slater determinants form a submanifold of the space of wave functions. This manifold is the *Grassmannian*, whose geometry is of central importance in mathematics, but still of few known applications in atomic and molecular physics or in theoretical chemistry, even though Slater determinants are key elements to electronic structure theory. This work shows how the geometry of the Grassmannian can be used for both theoretical considerations on the electron correlation and the practical optimization of a Slater determinant.

We have considered two approaches for an optimization process based on the Newton method. The first is using the standard parametrization by orbital rotations, and the second is an algorithm that explores the intrinsic geometry of the Grassmannian, as described by Absil *et al.* [39]. We showed that both algorithms are equivalent, in the sense that they lead to the same iterations (the same sequence of Slater determinants). However, the second algorithm allows a much more efficient implementation since it avoids the undesirable basis transformation step of the external wave function, that is not only time consuming, but destroys the original rank truncation of the wave function. Furthermore, the second algorithm uses directly the full coefficients matrix of the orbitals in the Slater determinant, that is a set of nonindependent parameters. It is often assumed that wave-function optimizations based on the Newton method cannot be performed in such condition, and

an independent set of parameters that covers the desired space of wave functions is necessary. Here, we showed that this is perfectly possible, as long as this is made carefully to project out the variations on the redundant parameters that do not change the wave function. This is done after considerations on the geometry of the underlying manifold [39].

The present algorithms converge quickly, typically in three iterations, as long that the starting point is reasonably close to the maximum overlap Slater determinant, such as the restricted Hartree-Fock Slater determinant for most of the systems. However, the procedures might converge to a relative maximum or to saddle points, if started with a poor initial guess. For such difficult cases, the present algorithms can be used in conjunction with the procedure of Zhang and Kollar [10], that converges more robustly, although in much more iterations.

Applications of the algorithm suggest that using the restricted Hartree-Fock wave function $|\Phi_{\text{HF}}\rangle$ to measure correlation is qualitatively equivalent to using $|\Phi_{\text{minD}}\rangle$, the Slater determinant that minimizes the distance to an external wave function $|\Psi_{\text{ext}}\rangle$. For most of the cases, $|\langle\Phi_{\text{minD}}|\Psi_{\text{ext}}\rangle|^2$ accounts for more than 99% of $|\langle\Phi_{\text{HF}}|\Psi_{\text{ext}}\rangle|^2$. However, there are quantitative differences when correlation is very large, and basis set truncation effect might be strong on both $|\langle\Phi_{\text{minD}}|\Psi_{\text{ext}}\rangle|$ and $|\langle\Phi_{\text{minD}}|\Phi_{\text{HF}}\rangle|$. We emphasize that the present examples are all singlet and closed-shell systems, with CISD wave functions based on a restricted Hartree-Fock reference, what forces $|\Phi_{\text{minD}}\rangle$ to be also spin restricted, and thus naturally close to $|\Phi_{\text{HF}}\rangle$. For instance, $|\Phi_{\text{minD}}\rangle$ is actually very far from the unrestricted Hartree-Fock wave function in cases of instabilities on the restricted Hartree-Fock wave function, as in the dissociation limit of H₂ and Li₂ molecules discussed here. Larger differences between $|\langle\Phi_{\text{minD}}|\Psi_{\text{CISD}}\rangle|$ and $|\langle\Phi_{\text{HF}}|\Psi_{\text{CISD}}\rangle|$ are thus expected for open-shell cases, where the spin restriction over $|\Phi_{\text{minD}}\rangle$ has to be relaxed. Furthermore, the single-reference CISD method is of limited

usage nowadays, and the evaluation of $|\langle \Phi_{\min D} | \Psi_{\text{CCSD}} \rangle|$, where $|\Psi_{\text{CCSD}}\rangle$ is the coupled cluster with single and double excitations [48,49], for instance, is more appealing. However, the full set of excited determinants would be needed for the present algorithms. One possible approximation is to consider only the projection of the CCSD, wave function into the space of up to doubly excited determinants, that is a CISD-like wave function. Hence, the present algorithms allow several numerical investigations on the quantification of electronic correlation.

ACKNOWLEDGMENTS

The authors thank the Dean's Office for Research of UFABC for providing the research facilities, and to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-Finance Code 001. Y.A.A. acknowledges the Grants No. 2017/21199-0 and No. 2018/04617-6, São Paulo Research Foundation (FAPESP). Y.A.A. acknowledges useful discussions with Prof. P. R. Taylor.

APPENDIX

1. Spatial orbitals and symmetry considerations

In this Appendix we present explicit formulas for the case where $|\Psi_{\text{ext}}\rangle$ is constructed from symmetry-adapted spatial orbitals, based on Abelian point groups. This means that the spin-orbital space [Eq. (5)] is given as the following direct sum of spaces:

$$\mathcal{W} = \mathcal{W}_1^\alpha \oplus \dots \oplus \mathcal{W}_g^\alpha \oplus \mathcal{W}_1^\beta \oplus \dots \oplus \mathcal{W}_g^\beta, \quad (\text{A1})$$

where the vector space associated to the irreducible representation (irrep) Γ and spin σ is of dimension M_Γ :

$$\mathcal{W}_\Gamma^\sigma = [\phi_1^\Gamma \otimes \sigma \wedge \dots \wedge \phi_{M_\Gamma}^\Gamma \otimes \sigma]. \quad (\text{A2})$$

In this condition, Eq. (23) becomes [28]

$$\hat{K} = \sum_\Gamma \sum_{(i,a) \in \Gamma} K_i^{a,\Gamma} (E_i^{a,\Gamma} - E_a^{i,\Gamma}), \quad (\text{A3})$$

where $E_q^{p,\Gamma} = a_{\alpha p, \Gamma}^\dagger a_{\alpha q, \Gamma} + a_{\beta p, \Gamma}^\dagger a_{\beta q, \Gamma}$ are the singlet excitation operators for the irrep Γ . Extension of the algorithm based in orbital rotations discussed in Sec. III A is straightforward, although care should be taken to the orbital ordering and the sign of coefficients.

a. Algorithm 2: Equations for a general $|\Psi_{\text{ext}}\rangle$

Function f_\diamond becomes

$$f_\diamond(U) = \sum_{\text{occ}(I) = \text{occ}(U)} C_I \prod_\Gamma F_{I_\alpha}^\Gamma F_{I_\beta}^\Gamma, \quad (\text{A4})$$

and Eq. (52) becomes

$$(\mathcal{H}_{\sigma\Gamma}^{\sigma'\Gamma'})_{qs}^{pr} ((\eta_\diamond U)_{\sigma'\Gamma'})_s^r = -(\mathcal{J}_{\sigma\Gamma})_q^p, \quad (\text{A5})$$

where $\sigma\Gamma$ indicates the block of the corresponding matrix associated to spin σ and irrep Γ . The notation “ $\text{occ}(I) = \text{occ}(U)$ ” indicates that only terms of $|\Psi_{\text{ext}}\rangle$ that have the same number of electrons as U in all blocks must be included. Indices p and r run over all orbitals of that symmetry

(M_Γ in number), whereas q and s run over the electrons in that spin and symmetry. In the following equations, quantities $\mathcal{J}_{\sigma\Gamma}$ and $\mathbf{G}_{I_\sigma}^\Gamma$ are two-index tensors of shape $(M_\Gamma, n_\sigma^\sigma)$, whereas $\mathcal{H}_{\sigma\Gamma}^{\sigma'\Gamma'}$ and $\tilde{\mathbf{H}}_{I_\sigma}^\Gamma$ are four-index quantities, of shape $(M_\Gamma, n_\sigma^\sigma, M_\Gamma, n_\sigma^{\sigma'})$ and $(M_\Gamma, n_\sigma^\sigma, M_\Gamma, n_\sigma^{\sigma'})$, respectively. The tensor product \otimes between a (M, n) quantity by a (M', n') quantity is the (M, n, M', n') quantity whose entries are

$$(A \otimes B)_{qs}^{pr} = A_q^p B_s^r. \quad (\text{A6})$$

$$\mathcal{J}_{\sigma\Gamma} = \Pi_{U_\sigma^\Gamma \perp} \sum_I C_I \left(\prod_{\{\sigma', \Gamma'\} \neq \{\sigma, \Gamma\}} F_{I_{\sigma'}}^{\Gamma'} \right) \mathbf{G}_{I_\sigma}^\Gamma, \quad (\text{A7})$$

$$\mathcal{H}_{\sigma\Gamma}^{\sigma'\Gamma'} = (\Pi_{U_\sigma^\Gamma \perp} \otimes \mathbb{1}) \sum_I C_I \left(\prod_{\{\sigma', \Gamma'\} \neq \{\sigma, \Gamma\}} F_{I_{\sigma'}}^{\Gamma'} \right) \tilde{\mathbf{H}}_{I_\sigma}^\Gamma, \quad (\text{A8})$$

$$\mathcal{H}_{\sigma\Gamma}^{\sigma''\Gamma''} = (\Pi_{U_\sigma^\Gamma \perp} \otimes \Pi_{U_{\sigma''}^{\Gamma''} \perp}) \sum_I C_I \left(\prod_{\substack{\{\sigma'', \Gamma''\} \neq \{\sigma', \Gamma'\} \\ \{\sigma'', \Gamma''\} \neq \{\sigma, \Gamma\}}} F_{I_{\sigma''}^{\Gamma''}} \right) \mathbf{G}_{I_\sigma}^\Gamma \otimes \mathbf{G}_{I_{\sigma''}^{\Gamma''}}^{\Gamma''}, \quad (\text{A9})$$

where the last equation holds for $\{\sigma, \Gamma\} \neq \{\sigma', \Gamma'\}$. These equations are obtained after considering the block-diagonal structure of the matrices $U|_I$, $(U \stackrel{q}{\leftarrow} e_p)|_I$, and $(U \stackrel{q}{\leftarrow} e_p \stackrel{s}{\leftarrow} e_r)|_I$, along with the fact that the determinant of a block-diagonal matrix is the product of the determinants of its blocks (see Sec. III B). If I has a different number of electrons than in U in any of its $\sigma\Gamma$ blocks [i.e., $\text{occ}(I) \neq \text{occ}(U)$] the matrices above have nonsquare blocks and their determinants are zero.

b. Algorithm 2: Equations for a restricted CISD wave function

Suppose now that $|\Psi_{\text{ext}}\rangle$ is a spin-restricted CISD wave function, based on a restricted and closed-shell reference Slater determinant. The reference determinant is given as

$$|\Phi_0\rangle = \phi_1^{\Gamma=1} \wedge \dots \wedge \phi_{n_g}^{\Gamma=g} \wedge \bar{\phi}_1^{\Gamma=1} \wedge \dots \wedge \bar{\phi}_{n_g}^{\Gamma=g} \\ = |\Phi_0\rangle_1 \wedge \dots \wedge |\Phi_0\rangle_g \wedge |\bar{\Phi}_0\rangle_1 \wedge \dots \wedge |\bar{\Phi}_0\rangle_g, \quad (\text{A10})$$

where, for example, the subspace of $[|\Phi_0\rangle]$ associated to alpha orbitals of irrep 1 is $[|\Phi_0\rangle_1]$ and so on. Overlines indicate beta spin. To simplify the notation, only the blocks where some excitation occurs will be shown, and the blocks not shown are assumed to be equal as in the reference determinant. Thus, for example,

$$|\Phi_i^a\rangle_\Gamma = |\Phi_0\rangle_1 \wedge \dots \wedge \phi_1^\Gamma \wedge \dots \wedge \phi_{i-1}^\Gamma \\ \wedge \phi_{i+1}^\Gamma \wedge \dots \wedge \phi_{n_r}^\Gamma \wedge \phi_a^\Gamma \wedge \dots \wedge |\bar{\Phi}_0\rangle_g \quad (\text{A11})$$

is an alpha single excitation from i to a in the irrep Γ . With this notation, the CISD wave function can be written as

$$\begin{aligned}
|\Psi_{\text{CISD}}\rangle &= C_0|\Phi_0\rangle + \sum_{\Gamma} \sum_{(i,a)\in\Gamma} C_i^{a,\Gamma} (|\Phi_i^a\rangle_{\Gamma} + |\bar{\Phi}_i^a\rangle_{\Gamma}) \\
&+ \sum_{\Gamma} \sum_{\substack{(i>j) \\ (a>b)}} \in\Gamma C_{ij}^{ab,\Gamma} (|\Phi_{ij}^{ab}\rangle_{\Gamma} + |\bar{\Phi}_{ij}^{ab}\rangle_{\Gamma}) \\
&+ \sum_{\Gamma} \sum_{\substack{(i,a) \\ (j,b)}} \in\Gamma C_{ij}^{ab,\Gamma} |\Phi_i^a\rangle_{\Gamma} \cdots |\Phi_j^b\rangle_{\Gamma} \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{A}_{ij}^{ab,\Gamma\Gamma'} (|\Phi_j^b\rangle_{\Gamma'} \cdots |\Phi_i^a\rangle_{\Gamma} \\
&\quad + |\bar{\Phi}_j^b\rangle_{\Gamma'} \cdots |\bar{\Phi}_i^a\rangle_{\Gamma}) \\
&+ \sum_{\Gamma>\Gamma'} \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{B}_{ij}^{ab,\Gamma\Gamma'} (|\Phi_j^b\rangle_{\Gamma'} \cdots |\bar{\Phi}_i^a\rangle_{\Gamma} \\
&\quad + |\Phi_i^a\rangle_{\Gamma} \cdots |\bar{\Phi}_j^b\rangle_{\Gamma'}) \\
&+ \sum_{\substack{I \text{ doubles over } |\Phi_0\rangle \\ \text{occ}(I) \neq \text{occ}(0)}} C_I |\Phi_I\rangle, \tag{A12}
\end{aligned}$$

where $C_{ij}^{ab,\Gamma} = C_{ji}^{ba,\Gamma}$, as it is a restricted wave function. As will be seen below, $\mathcal{A}_{ij}^{ab,\Gamma\Gamma'}$ and $\mathcal{B}_{ij}^{ab,\Gamma\Gamma'}$ always appear summed, and thus the coefficients of double excitations arising as product of single excitations at different blocks are merged in a single quantity \mathcal{D} :

$$\mathcal{D}_{ij}^{ab,\Gamma\Gamma} = C_{ij}^{ab,\Gamma}, \tag{A13}$$

$$\mathcal{D}_{ij}^{ab,\Gamma\Gamma'} = \mathcal{A}_{ij}^{ab,\Gamma\Gamma'} + \mathcal{B}_{ij}^{ab,\Gamma\Gamma'} \quad \text{for } \Gamma \neq \Gamma'. \tag{A14}$$

Determinants that have some spin or irrep with a number of electrons different than in the reference determinant are collected in the last term, and they contribute neither to f_{\diamond} nor to the matrices used in the optimization. Applying this wave function in the equations of Appendix A 1 a, the following equations are obtained. At first we define some intermediates:

$$\mathcal{F}_0 = \prod_{\Gamma} F_{I_0^{\Gamma}}^2, \tag{A15}$$

$$\mathcal{F}_0^{\Gamma} = \prod_{\Gamma' \neq \Gamma} F_{I_0^{\Gamma'}}^2, \tag{A16}$$

$$\mathcal{F}_0^{\Gamma\Gamma'} = \prod_{\substack{\Gamma'' \neq \Gamma \\ \Gamma'' \neq \Gamma'}} F_{I_0^{\Gamma''}}^2, \tag{A17}$$

with analogous definitions for $\mathcal{F}_0^{\Gamma\Gamma'\Gamma''}$ and $\mathcal{F}_0^{\Gamma\Gamma'\Gamma''\Gamma''''}$;

$$\begin{aligned}
\mathcal{K}^{\Gamma\Gamma'} &= 2 \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'} F_{I_0^{\Gamma}} F_{I_0^{\Gamma'}} F_{I_i^{a,\Gamma}} F_{I_j^{b,\Gamma'}} \\
&\text{for } \Gamma \neq \Gamma'. \tag{A18}
\end{aligned}$$

In these equations, I_0^{Γ} is the multi-index of the reference for irrep Γ , whereas $I_i^{a,\Gamma}$ is the multi-index for the single excitation from i to a , also in Γ . The other type of multi-index that appears is $I_{ij}^{ab,\Gamma}$, for double excitations at irrep Γ :

$$\begin{aligned}
\mathcal{L}^{\Gamma} &= 2F_{I_0^{\Gamma}} \left(\sum_{(i,a)\in\Gamma} C_i^{a,\Gamma} F_{I_i^{a,\Gamma}} + \sum_{\substack{(i>j) \\ (a>b)}} \in\Gamma C_{ij}^{ab,\Gamma} F_{I_{ij}^{ab,\Gamma}} \right) \\
&+ \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}} F_{I_j^{b,\Gamma}}. \tag{A19}
\end{aligned}$$

The factor 2 takes into account the contributions from alpha and beta excitations, both totally within the same irrep Γ . With the so far defined quantities, we are able to calculate $f_{\diamond}(U)$, assuming U orthonormal:

$$f_{\diamond}(U) = C_0 \mathcal{F}_0 + \sum_{\Gamma} \mathcal{F}_0^{\Gamma} \mathcal{L}^{\Gamma} + \sum_{\Gamma>\Gamma'} \mathcal{F}_0^{\Gamma\Gamma'} \mathcal{K}^{\Gamma\Gamma'}. \tag{A20}$$

It is not difficult to see the origin of each term of Eq. (A20) [compare to Eq. (A4)]: the first is the contribution from the reference determinant, that is the product of all $F_{I_0^{\Gamma}}$, for all Γ and for each spin. Since it is a closed-shell restricted wave function, this is just \mathcal{F}_0 . The second term of Eq. (A20) is the contribution of all excitations within the same irrep: for all such excitations, the other irrep blocks contribute with a $F_{I_0^{\Gamma'}}^2$, that form a common \mathcal{F}_0^{Γ} ; the contribution of the irrep in question is the CI coefficient, times an appropriate $F_{I_{ij}^{ab,\Gamma}}$, as can be seen in Eq. (A19). The last term is the contribution of excitations at mixed irreps (say Γ and Γ'): each determinant contributes with one $F_{I_0^{\Gamma}}$, one $F_{I_0^{\Gamma'}}$ (from the spins where no excitations occurred, whichever they are), and the $F_{I_i^{a,\Gamma}}$ and $F_{I_j^{b,\Gamma'}}$ of corresponding single excitations. This is clearly seen in (A18), and the contribution of remaining irreps forms $\mathcal{F}_0^{\Gamma\Gamma'}$.

For the matrices \mathcal{H} and \mathcal{J} , Eqs. (A7) to (A9), the interpretation is similar. We define the intermediates [the quantities $\mathbf{G}_{I_i^{a,\Gamma}}$, $\hat{\mathbf{G}}_{I_i^a}^{\Gamma}$, and \mathcal{M}^{Γ} have shape (M_{Γ}, n_{Γ}) , whereas \mathcal{G}^{Γ} , $\hat{\mathbf{H}}$, and \mathcal{H}^{Γ} have shape $(M_{\Gamma}, n_{\Gamma}, M_{\Gamma}, n_{\Gamma})$]:

$$\hat{\mathbf{G}}_{I_i^a}^{\Gamma} = F_{I_0^{\Gamma}} \mathbf{G}_{I_i^{a,\Gamma}} + F_{I_i^{a,\Gamma}} \mathbf{G}_{I_0^{\Gamma}}, \tag{A21}$$

$$\bar{C}_i^{a,\Gamma} = C_i^{a,\Gamma} + \sum_{\bar{\Gamma} \neq \Gamma} \frac{1}{F_{I_0^{\bar{\Gamma}}}} \sum_{(j,b)\in\bar{\Gamma}} F_{I_j^{b,\bar{\Gamma}}} \mathcal{D}_{ij}^{ab,\Gamma\bar{\Gamma}}, \tag{A22}$$

$$\mathcal{M}^{\Gamma} = C_0 F_{I_0^{\Gamma}} \mathbf{G}_{I_0^{\Gamma}} + \sum_{(i,a)\in\Gamma} \bar{C}_i^{a,\Gamma} \hat{\mathbf{G}}_{I_i^a}^{\Gamma}$$

$$\begin{aligned}
&+ \sum_{\substack{(i>j) \\ (a>b)}} \in\Gamma C_{ij}^{ab,\Gamma} (F_{I_0^{\Gamma}} \mathbf{G}_{I_{ij}^{ab,\Gamma}} + F_{I_{ij}^{ab,\Gamma}} \mathbf{G}_{I_0^{\Gamma}}) \\
&+ \sum_{\substack{(i,a)\in\Gamma \\ (j,b)\in\Gamma}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}} \mathbf{G}_{I_j^{b,\Gamma}}, \tag{A23}
\end{aligned}$$

$$\begin{aligned}
 \mathcal{G}^\Gamma &= C_0 \mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_0^\Gamma} \\
 &+ \sum_{(i,a) \in \Gamma} \bar{C}_i^{a,\Gamma} (\mathbf{G}_{I_i^{a,\Gamma}} \otimes \mathbf{G}_{I_0^\Gamma} + \mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_i^{a,\Gamma}}) \\
 &+ \sum_{\substack{(i>j) \\ (a>b)}} \{ \in \Gamma \} C_{ij}^{ab,\Gamma} (\mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_{ij}^{ab,\Gamma}} + \mathbf{G}_{I_{ij}^{ab,\Gamma}} \otimes \mathbf{G}_{I_0^\Gamma}) \\
 &+ \sum_{\substack{(i,a) \\ (j,b)}} \{ \in \Gamma \} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} \mathbf{G}_{I_i^{a,\Gamma}} \otimes \mathbf{G}_{I_j^{b,\Gamma}}, \quad (\text{A24})
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{H}^\Gamma &= C_0 F_{I_0^\Gamma} \tilde{\mathbf{H}}_{I_0^\Gamma} + \sum_{(i,a) \in \Gamma} \bar{C}_i^{a,\Gamma} (F_{I_0^\Gamma} \tilde{\mathbf{H}}_{I_i^{a,\Gamma}} + F_{I_i^{a,\Gamma}} \tilde{\mathbf{H}}_{I_0^\Gamma}) \\
 &+ \sum_{\substack{(i>j) \\ (a>b)}} \{ \in \Gamma \} C_{ij}^{ab,\Gamma} (F_{I_0^\Gamma} \tilde{\mathbf{H}}_{I_{ij}^{ab,\Gamma}} + F_{I_{ij}^{ab,\Gamma}} \tilde{\mathbf{H}}_{I_0^\Gamma}) \\
 &+ \sum_{\substack{(i,a) \\ (j,b)}} \{ \in \Gamma \} \mathcal{D}_{ij}^{ab,\Gamma\Gamma} F_{I_i^{a,\Gamma}} \tilde{\mathbf{H}}_{I_j^{b,\Gamma}}. \quad (\text{A25})
 \end{aligned}$$

Note that the terms from ‘‘single excitations’’ in these quantities contain a contribution from double excitations that occur partially in Γ and partially in another irrep [see Eq. (A22)], and thus represent indeed single excitations from the point of view of irrep Γ .

Finally,

$$\mathcal{J}_\Gamma = \Pi_{U^\Gamma \perp} \left\{ \mathcal{F}_0^\Gamma \mathcal{M}^\Gamma + \left(\sum_{\substack{\bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} > \Gamma \\ \bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} \neq \Gamma}} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{L}^{\bar{\Gamma}} + \sum_{\substack{\bar{\Gamma} > \Gamma \\ \bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} \neq \Gamma}} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{K}^{\bar{\Gamma}\bar{\Gamma}} \right) F_{I_0^\Gamma} \mathbf{G}_{I_0^\Gamma} \right\}, \quad (\text{A26})$$

$$\mathcal{H}_\Gamma^\Gamma = (\Pi_{U^\Gamma \perp} \otimes \mathbb{1}) \left\{ \mathcal{F}_0^\Gamma (\mathcal{H}^\Gamma + (\mathbb{1} \otimes \Pi_{U^\Gamma \perp}) \mathcal{G}^\Gamma) + \left(\sum_{\substack{\bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} > \Gamma \\ \bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} \neq \Gamma}} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{L}^{\bar{\Gamma}} + \sum_{\substack{\bar{\Gamma} > \Gamma \\ \bar{\Gamma} \neq \Gamma \\ \bar{\Gamma} \neq \Gamma}} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{K}^{\bar{\Gamma}\bar{\Gamma}} \right) (F_{I_0^\Gamma} \tilde{\mathbf{H}}_{I_0^\Gamma} + \mathbf{G}_{I_0^\Gamma} \otimes (\Pi_{U^\Gamma \perp} \mathbf{G}_{I_0^\Gamma})) \right\}, \quad (\text{A27})$$

$$\begin{aligned}
 \mathcal{H}_{\Gamma'}^{\Gamma'} &= 2(\Pi_{U^{\Gamma'} \perp} \otimes \Pi_{U^{\Gamma'} \perp}) \left\{ \mathcal{F}_0^{\Gamma\Gamma'} \left\{ F_{I_0^{\Gamma'}} \mathcal{M}^\Gamma \otimes \mathbf{G}_{I_0^{\Gamma'}} + F_{I_0^\Gamma} \mathbf{G}_{I_0^\Gamma} \otimes \mathcal{M}^{\Gamma'} + \frac{1}{2} \sum_{\substack{(i,a) \in \Gamma \\ (j,b) \in \Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \hat{\mathbf{G}}_{ia}^\Gamma \otimes \hat{\mathbf{G}}_{jb}^{\Gamma'} \right. \right. \\
 &- \left. \left(\sum_{(i,a) \in \Gamma} \hat{\mathbf{G}}_{ia}^\Gamma \sum_{(j,b) \in \Gamma'} F_{I_j^{b,\Gamma'}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \right) \otimes \mathbf{G}_{I_0^{\Gamma'}} - \mathbf{G}_{I_0^\Gamma} \otimes \left(\sum_{(j,b) \in \Gamma'} \hat{\mathbf{G}}_{jb}^{\Gamma'} \sum_{(i,a) \in \Gamma} F_{I_i^{a,\Gamma}} \mathcal{D}_{ij}^{ab,\Gamma\Gamma'} \right) \right\} \\
 &+ \left. \left(-C_0 \mathcal{F}_0^{\Gamma\Gamma'} + \sum_{\bar{\Gamma} \neq \Gamma, \Gamma'} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{L}^{\bar{\Gamma}} + \sum_{\substack{\bar{\Gamma} > \Gamma \\ \bar{\Gamma} \neq \Gamma, \Gamma' \\ \bar{\Gamma} \neq \Gamma, \Gamma'}} \mathcal{F}_0^{\Gamma\bar{\Gamma}} \mathcal{K}^{\bar{\Gamma}\bar{\Gamma}} \right) F_{I_0^\Gamma} F_{I_0^{\Gamma'}} \mathbf{G}_{I_0^\Gamma} \otimes \mathbf{G}_{I_0^{\Gamma'}} \right\}. \quad (\text{A28})
 \end{aligned}$$

By inspection on these expressions one can see that the quantities F_I and \mathbf{G}_I for single excitations, $F_{I_i^{a,\Gamma}}$ and $\mathbf{G}_{I_i^{a,\Gamma}}$, are used often and their storage does not pose a problem. Other quantities such as $\tilde{\mathbf{H}}_I$ and those associated to double excitations are too many for storage, but they are used only once. Thus, an efficient implementation of these equations can be made that explores these

facts. Furthermore, the following relations among F_I , \mathbf{G}_I , and \mathbf{H}_I hold, and can be used to calculate some of these quantities from others:

$$F_{I_i^a} = (-1)^{i+n} \sum_{q=1}^n U_q^a (\mathbf{G}_{I_0})_q^i, \quad (\text{A29})$$

$$F_{I_{ij}^{ab}} = (-1)^{i+n+(b>a)} \sum_{q=1}^n U_q^a (\mathbf{G}_{I_j^b})_q^i, \quad (\text{A30})$$

$$(\mathbf{G}_{I_i^a})_q^p = \begin{cases} (-1)^{i+n} \sum_{q'=1}^n (1 - \delta_{qq'}) U_{q'}^a (\mathbf{H}_{I_0})_{qq'}^{pi}, & p \neq a \\ (-1)^{i+n} (\mathbf{G}_{I_0})_q^i, & p = a \end{cases} \quad (\text{A31})$$

where $(b > a) = 1$ if $b > a$, 0 otherwise.

-
- [1] P. Gersdorf, W. John, J. P. Perdew, and P. Ziesche, *Int. J. Quantum Chem.* **61**, 935 (1997).
- [2] C. L. Benavides-Riveros, N. N. Lathiotakis, C. Schilling, and M. A. L. Marques, *Phys. Rev. A* **95**, 032507 (2017).
- [3] L. Ding and C. Schilling, *J. Chem. Theory Comput.* **16**, 4159 (2020).
- [4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
- [5] J. M. Myers and T. T. Wu, *Quantum Inf. Proc.* **9**, 239 (2010).
- [6] D. Bures, *Trans. Am. Math. Soc.* **135**, 199 (1969).
- [7] M. Hübner, *Phys. Lett. A* **163**, 239 (1992).
- [8] I. D'Amico, J. P. Coe, V. V. Franca, and K. Capelle, *Phys. Rev. Lett.* **106**, 050401 (2011).
- [9] C. L. Benavides-Riveros, N. N. Lathiotakis, and M. A. L. Marques, *Phys. Chem. Chem. Phys.* **19**, 12655 (2017).
- [10] J. M. Zhang and M. Kollar, *Phys. Rev. A* **89**, 012504 (2014).
- [11] J. M. Zhang and N. J. Mauser, *Phys. Rev. A* **94**, 032513 (2016).
- [12] A. A. Borisenko and Y. A. Nikolaevski, *Russ. Math. Surveys* **46**, 45 (1991).
- [13] Đ. Baralić, *Teach. Math.* **XIV**, 147 (2011).
- [14] P. Griffiths and J. Harris, *Principles of Algebraic Geometry* (Wiley, New York, 1978).
- [15] W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry*, Vol. 1 (Cambridge University Press, Cambridge, 1994).
- [16] D. J. Rowe, A. Ryman, and G. Rosensteel, *Phys. Rev. A* **22**, 2362 (1980).
- [17] P. Cassam-Chenaï, *J. Math. Chem.* **15**, 303 (1994).
- [18] A. I. Panin, [arXiv:0707.3796](https://arxiv.org/abs/0707.3796).
- [19] E. Chiumiento and M. Melgaard, *J. Geom. Phys.* **62**, 1866 (2012).
- [20] É. Polack, A. Mikhalev, G. Dusson, B. Stamm, and F. Lipparini, *Mol. Phys.* **118**, e1779834 (2020).
- [21] J. R. Ruiz-Tolosa and E. Castillo, *From Vectors to Tensors* (Springer, Berlin, 2005).
- [22] R. M. Bowen and C. C. Wang, *Introduction to Vectors and Tensors* (Dover, New York, 2008).
- [23] K. C. Mundim, *J. Phys. (France)* **50**, 11 (1989).
- [24] K. C. Mundim and M. S. P. Mundim, *Rev. Bras. Ens. Fis.* **19**, 209 (1997).
- [25] A. Vourdas, *J. Phys. A: Math. Theor.* **51**, 445301 (2018).
- [26] W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry*, Vol. 2 (Cambridge University Press, Cambridge, 1994).
- [27] D. J. Thouless, *Nucl. Phys.* **21**, 225 (1960).
- [28] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, New York, 2000).
- [29] J. Linderberg and Y. Öhrn, *Int. J. Quantum Chem.* **XII**, 161 (1977).
- [30] E. Dalgaard and P. Jørgensen, *J. Chem. Phys.* **69**, 3833 (1978).
- [31] D. L. Yeager and P. Jørgensen, *J. Chem. Phys.* **71**, 755 (1979).
- [32] P. E. M. Siegbahn, A. Heiberg, B. O. Roos, and B. Levy, *Phys. Scr.* **21**, 323 (1980).
- [33] P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, *J. Chem. Phys.* **74**, 2384 (1981).
- [34] H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **82**, 5053 (1985).
- [35] H.-J. Werner, Matrix-formulated direct multiconfiguration self-consistent field and multiconfiguration reference configuration-interaction methods, in *Advances in Chemical Physics* (Wiley, New York, 1987), pp. 1–62.
- [36] R. Shepard, *The Multiconfiguration Self-consistent Field Method* (Wiley, New York, 1987), pp. 63–200.
- [37] B. O. Roos, *The Complete Active Space Self-consistent Field Method and its Applications in Electronic Structure Calculations* (Wiley, New York, 1987), pp. 399–445.
- [38] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevA.102.052803> for an animation that illustrates the concepts discussed in Sec. II B.
- [39] P. A. Absil, R. Mahony, and R. Sepulchre, *Acta Appl. Math.* **80**, 199 (2004).
- [40] P. Å. Malmqvist, *Int. J. Quantum Chem.* **XXX**, 479 (1986).
- [41] O. Travis E, *A guide to NumPy* (Trelgol Publishing, Austin, 2006).
- [42] S. van der Walt, S. C. Colbert, and G. Varoquaux, *Comput. Sci. Eng.* **13**, 22 (2011).
- [43] W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).
- [44] J. Schliemann, J. I. Cirac, M. Kuś, M. Lewenstein, and D. Loss, *Phys. Rev. A* **64**, 022303 (2001).
- [45] MOLPRO, version 2019 a package of *ab initio* programs. H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, and others, see <http://www.molpro.net>.
- [46] W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- [47] T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- [48] G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- [49] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).