



**HAL**  
open science

## 2D-Block Geminals: a non 1-orthogonal and non 0-seniority model with reduced computational complexity

Patrick Cassam-Chenaï, Thomas Perez, Davide Accomasso

### ► To cite this version:

Patrick Cassam-Chenaï, Thomas Perez, Davide Accomasso. 2D-Block Geminals: a non 1-orthogonal and non 0-seniority model with reduced computational complexity. 2022. hal-03763128v3

**HAL Id: hal-03763128**

**<https://hal.science/hal-03763128v3>**

Preprint submitted on 5 Nov 2022 (v3), last revised 3 Jan 2023 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## 2D-Block Geminals: a non 1-orthogonal and non 0-seniority model with reduced computational complexity

Patrick Cassam-Chenai, Thomas Perez<sup>1</sup> and Davide Accomasso<sup>2</sup>

<sup>1</sup>*Université Côte d'Azur, LJAD, UMR 7351, 06100 Nice, France*

<sup>2</sup>*Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via Moruzzi 13, 56124 Pisa, Italy*

We present a new geminal product wave function ansatz where the geminals are not constrained to be strongly orthogonal nor to be of seniority-zero. Instead, we introduce weaker orthogonality constraints between geminals which significantly lower the computational effort, without sacrificing the indistinguishability of the electrons. That is to say, the electron pairs corresponding to the geminals are not fully distinguishable, and their product has still to be antisymmetrized according to the Pauli principle to form a *bona fide* electronic wave function. Our geometrical constraints translate into simple equations involving the traces of products of our geminal matrices. In the simplest non-trivial model, a set of solutions is given by block-diagonal matrices where each block is of size 2x2 and consists of either a Pauli matrix or a normalized diagonal matrix, multiplied by a complex parameter to be optimized. With this simplified ansatz for geminals, the number of terms in the calculation of the matrix elements of quantum observables is considerably reduced. A proof of principle is reported and confirms that the ansatz is more accurate than strongly orthogonal geminal products while remaining computationally affordable.

Keywords:

Strongly correlated electron pairs, antisymmetric geminal products, Pauli matrices

Suggested running head:

2D-Block Geminals

All correspondance to be send to P. Cassam-Chenai,

cassam@unice.fr,

## I. INTRODUCTION

The accurate, *ab initio* calculation of the electronic energy levels of a molecular system is still a challenging problem with exponential complexity with respect to the number of particles. This is particularly true for the so-called “strongly correlated” systems, where many Slater determinants have to be included in the wave function to achieve a qualitatively correct description of the molecular state. In such systems, it has been established that better ansätze can be based on an electron pair model, that is to say, an antisymmetric product of two-electron wave functions, called “geminals”.

However, without further restrictions, such a model has still a factorial computational cost with respect to the number of electronic orbitals and its applicability is therefore limited to small systems. Although an exponential scaling seems inherent in general fermionic quantum systems<sup>1</sup>, there is hope to achieve a better scaling for special classes of Fermionic systems. Reducing the scaling of geminal methods’ computational cost to a polynomial one is an active research field<sup>2-5</sup>. We postpone to the next section a review of the most popular proposals. Here, we just note that they often enforce the “strong orthogonality” constraint which breaks the indistinguishability of electron pairs, and is thus at odds with the Pauli principle, or in non-orthogonal cases, they have been limited to seniority-zero wave functions with very few exceptions.

In the past, we have introduced generalized orthogonality constraints between geminals to reduce the computational effort, without sacrificing the indistinguishability of the electrons nor assuming a given seniority<sup>6,7</sup>. In other words, the electron pairs corresponding to the geminals were not fully distinguishable, so that geminal products had still to be antisymmetrized according to the Pauli principle to form *bona fide* electronic wave functions, and each geminal combined linearly electronic configurations of mixed seniorities. The aim of the present work is to pursue our exploration of orthogonality constraints to improve the computational affordability of antisymmetric product of geminals (APG).

Recall that an APG wave function can be described in an alternative and practical way, by a set of matrices, one for each geminal. Our geometrical constraints translated in terms of these geminal matrices, into simple equations involving the traces of their products. Here, we will impose further geometrical constraints whose solutions will lead us to consider block-diagonal matrices, where each block is at most of size  $2 \times 2$ . These blocks are essentially Pauli or diagonal matrices, multiplied by a complex parameter to be optimized. With this simple ansatz for geminals, the number of terms in the calculation of quantum observable matrix elements is considerably reduced.

The paper is organized as follows: In the next section we present an overview of existing APG-based methods. Then, we obtain an overlap formula for the general APG case, which exhibits its full complexity and serves us as a starting point to elaborate new constraints able to make it more amenable to practical computations. Next, we propose to enforce the so-called “permutationally invariant 2-orthogonality constraints” and show how we arrive at our 2D-block geminal model. Finally, we obtain the necessary quantities to implement our ansatz and provide a proof of principle that it is able to give strictly lower energies than strongly orthogonal geminal products. In the last section, we conclude on the prospect of our method.

## II. ANTISYMMETRIZED PRODUCT OF GEMINALS (APG) ANSÄTZE: AN OVERVIEW

Many APG models have appeared in the literature. We provide here a quick overview to help the reader in situating the model proposed in the following sections. We recall in Appendix A, the generalized concept of seniority, and the definition of  $p$ -orthogonality, because these two notions make the constraints generally imposed on APG wave functions amenable to a synthetic presentation.

## A. General APG ansatz

A general APG can be written as

$$\Phi = g^1 \wedge g^2 \wedge \cdots \wedge g^p \quad (1)$$

with

$$g^r = \sum_{i=1}^{2m-1} \sum_{j>i}^{2m} \lambda_{i,j}^r \varphi_i \wedge \varphi_j \quad . \quad (2)$$

The  $\varphi_i$ 's form a basis of one-electron wavefunctions, not necessarily spin-adapted, i.e., they can be linear combinations involving both spin  $\alpha$  and spin  $\beta$  spin-orbitals. Each geminal  $g^r$  is parametrized by  $m(2m - 1)$  independent scalar (real or complex) numbers  $\lambda_{i,j}^r$ .

However, even in the spin-adapted case, optimizing such ansatz is extremely costly from a computational standpoint<sup>8</sup> and has rarely been attempted<sup>7,9,10</sup>. All popular geminal ansätze can be seen as APG with additional constraints. The latter can be divided into “intra-geminal” constraints, bearing on the spin-orbitals of the geminal expansion Eq.(2), and “inter-geminal” constraints, imposing certain relationships between the geminals of the product Eq.(1).

## B. Intra-geminal constraints

### 1. *APsetG ansatz:*

A first approach consists in partitioning the spin-orbital basis into two subsets of cardinal  $m$ ,  $\mathcal{H} = \mathcal{H}_\alpha \oplus \mathcal{H}_\beta$ , ( $\dim \mathcal{H}_\alpha = \dim \mathcal{H}_\beta = m$ ), and to expand the geminals on two-electron basis functions of generalized seniority equal to 2 with respect to this partition:

$$g_{\text{set}}^r = \sum_{i=1}^m \sum_{j=1}^m \lambda_{i,j}^r \varphi_{2i-1} \wedge \varphi_{2j} \quad , \quad (3)$$

where the subscript “set” hints to the fact that we have two disjoint sets of identical cardinal. One can choose for example to gather all spin-orbitals of spin  $\alpha$  in the first set and all those of spin  $\beta$  in the second set. Then, one has wave functions that are eigenfunctions of  $\hat{S}_z$  associated with the eigenvalue 0.

Geminals of this type are parametrized by  $m^2$  independent scalar numbers.

The Grassmann product of those geminals is called an APsetG (“Antisymmetric Product of set-divided Geminals”)<sup>11</sup> wave function:

$$\Phi_{\text{APsetG}} = \bigwedge_{r=1}^k g_{\text{set}}^r \quad . \quad (4)$$

## 2. *APIG ansatz:*

An alternative idea is to partition the spin-orbital basis set into  $m$  shells of the form  $(\varphi_{2i-1}, \varphi_{2i})$  and to build geminals out of seniority 0 two-electron functions exclusively:

$$g_{\text{I}}^r = \sum_{i=1}^m \lambda_i^r \varphi_{2i-1} \wedge \varphi_{2i} \quad , \quad (5)$$

where the subscript “I” stands for “Interacting”, as the Grassmann product of those geminals is known as the APIG wave function (“Antisymmetric Product of Interacting Geminals”)<sup>11</sup>

$$\Phi_{\text{APIG}} = \bigwedge_{r=1}^k g_{\text{sen0}}^r \quad . \quad (6)$$

Each geminals in the product is parametrized by  $m$  independent scalar numbers.

Note that, if the shells of the partition are such that, for all  $i$ ,  $\varphi_{2i-1}$  and  $\varphi_{2i}$  are spin-orbitals of the same spatial part and of opposite spin, one retrieve the usual concept

of seniority, and the APIG wave functions are those considered by Silver<sup>12</sup> and revived recently by Limacher<sup>13</sup>. In this case a more natural notation is:

$$g_I^r = \sum_{i=1}^m \lambda_i^r \varphi_i \wedge \bar{\varphi}_i \quad . \quad (7)$$

Most of other APG ansätze in use are particular cases of the APIG ansatz, hence of seniority 0. Let us mention the antisymmetric product of rank-two geminals, APr2G, where the coefficients are constrained to assume the form<sup>14,15</sup>:  $\lambda_i^r = \frac{1}{a_r \epsilon_i - b_i \lambda_r}$ ,  $a_r, \epsilon_i, b_i, \lambda_r \in \mathbb{C}$ . When these numbers are derived as eigenstate coefficients of reduced Bardeen-Cooper-Schrieffer (BCS) model Hamiltonians<sup>16</sup>, one obtains the so-called Richardson–Gaudin (RG) geminals which are currently under active development.<sup>2,17</sup>

## C. Inter-geminal constraints

### 1. *AP1roG ansatz*:

Another important subcase of the APIG model consists in imposing the following additional constraint: each geminal is only allowed to have a non-zero coefficient on a single pair of spin-orbitals (numbered from 1 to  $k$ ) pertaining to an Hartree-Fock type reference wave function (with  $k$  doubly occupied orbitals):

$$g_{1ro}^r = \varphi_r \wedge \bar{\varphi}_r + \sum_{i=k+1}^m \lambda_i^r \varphi_i \wedge \bar{\varphi}_i \quad , \quad (8)$$

where the subscript “1ro” stands for “1-reference orbital”. Note that the constraint has an intra-geminal character but has also an inter-geminal one, since two different geminals must have different doubly occupied reference orbitals. Each 1ro-geminal is parametrized by  $m - k$  independent scalar numbers.



The Grassmann product of those geminals is called the “AP1roG” wave function (“Antisymmetric Product of 1-reference orbital Geminals”)<sup>15,18–22</sup>:

$$\Phi_{\text{AP1roG}} = \bigwedge_{r=1}^k g_{1\text{ro}}^r \quad . \quad (9)$$

### 2. *APSG ansatz:*

The 1-orthogonality constraint (see Appendix VII B) between the geminals in the Grassmann product has often been added as an inter-geminal condition. One obtains in this way an APSG (“Antisymmetrized Product of Strongly-orthogonal Geminals”) wave functions, which have been extensively studied<sup>12,23–25,27–34</sup>. This ansatz, as well as its subcase, the generalized valence bond perfect-pairing (GVB-PP) wave function<sup>35,36</sup>, have also been investigated for the calculation of excited states, with a time-dependent linear response theory approach<sup>37</sup>.

The 1-orthogonality constraint reduces drastically the computational cost, however, it is usually too stringent to achieve chemical accuracy.

### 3. *AGP ansatz:*

At the other extreme, one can take all  $k$  geminals, not just non-orthogonal but even equal. Then, one obtains a product wave function called “AGP” for “Antisymmetrized Geminal Power”<sup>38</sup>. Dropping the subscript and denoting  $g$  the geminal involved, we have:

$$\Phi_{\text{AGP}} = g^{\wedge k} = \underbrace{g \wedge \dots \wedge g}_{k \text{ times}} \quad . \quad (10)$$

#### 4. *GMFCI p-orthogonal ansätze:*

The  $p$ -orthogonality constraint,  $1 \leq p \leq \inf(n_1, n_2)$  between an  $n_1$ -electron wave function and an  $n_2$ -electron one (see Appendix VII B), can be tuned from strong orthogonality ( $p = 1$ ), to “weak orthogonality” ( $p = \inf(n_1, n_2)$ ), to restrict variational freedom. Note that, in the case ( $n_1 = n_2 = n$ ), weak orthogonality ( $p = n$ ) is just the usual orthogonality between  $n$ -electron wave functions.

Two such constraints, in between the APSG and AGP extreme cases, have been considered in the past for Geminal Mean Field Configuration Interaction (GMFCI) wave functions<sup>7</sup>: (i) 2-orthogonality between every pairs of geminals ( $g^{r_1}, g^{r_2}$ ) of the APG; (ii) 2-orthogonality between any geminal  $g^r$  and any Grassmann product of the remaining ones  $\bigwedge_{s \neq r} g^s$ .

#### 5. *Various post-treatments:*

There have been many proposals to improve APSG or AP1roG wave functions by adding perturbative corrections<sup>39–45</sup>. Other attempts have consisted in applying linearized coupled cluster corrections to APSG or AP1roG reference wave functions<sup>46,47</sup>. However, these improvements were at the expense of the simplicity of the antisymmetrized geminal product model.

Many recent works aim at improving the AGP ansatz<sup>56,57</sup>. A promising method called “CJAGP” for “Cluster-Jastrow Antisymmetrized Geminal Power”<sup>48–50</sup>, or its “JAGP” particular case when atomic orbitals are not optimized<sup>51–53</sup>, have been proposed by Neuscamman and can achieve polynomial cost, but with a large prefactor. Beside the latter approach qualified as “subtractive” according to Neuscamman’s interesting terminology, AGPs were also used as reference wave functions for more classical “additive” approaches<sup>54,55</sup>.

Note also that, very recently, the AGP ansatz<sup>58</sup> and an unitary version of AP1roG<sup>59</sup> have been ported to quantum computers.

In order to better understand the computational cost reductions occurring in the main ansätze just reviewed, and to propose new constrained APG models, we find appropriate to return first to the general case in the next section.

### III. GENERAL APG OVERLAP FORMULA

In general, in quantum chemistry, the observables of interest, such as the Hamiltonian, contain at most two-particle operators. So, the computational cost of an electronic wave function method is essentially given by the cost of calculating overlap integrals.

In this section, we provide a general formula, proved in Appendix VII C, for the overlap between two wave functions which are products of geminals, i.e. to compute  $\langle \Phi_1 \wedge \dots \wedge \Phi_n | \Phi'_1 \wedge \dots \wedge \Phi'_n \rangle$ , where the  $\Phi_k$ 's and the  $\Phi'_k$ 's are general geminals with  $S_z = 0$ .

We will use a spin-adapted orthonormal basis set for the one-electron Hilbert space  $\mathcal{H}$  (of finite dimension  $2m$ ), that is made of two subsets  $(\varphi_i)_{1 \leq i \leq m}$  and  $(\bar{\varphi}_i)_{1 \leq i \leq m}$ , the latter being, respectively, spin  $\alpha$  and spin  $\beta$  spin-orbitals of the same set of orbitals.

We associate with each geminal,  $\Phi_k$ , a matrix of size  $m \times m$ , denoted  $C_k$ , defined as follows:

$$\forall k \in \{1, \dots, n\}, C_k = \left( \langle \Phi_k | \varphi_i \wedge \bar{\varphi}_j \rangle \right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq m}}, \quad (11)$$

so that:

$$\forall k \in \{1, \dots, n\}, \Phi_k = \sum_{1 \leq i, j \leq m} (C_k)_{i,j} \varphi_i \wedge \bar{\varphi}_j \quad . \quad (12)$$

Then the scalar product of two geminals can be expressed as the trace of the product of two matrices:

$$\langle \Phi_i | \Phi_j \rangle = \text{tr}(C_i^\dagger C_j) \quad . \quad (13)$$

Note that, the matrix associated with a singlet geminal, will be symmetric (even if we work with complex numbers), and that of a triplet geminal, will be skew-symmetric.

Let  $\Phi'_k$ 's be geminals whose associated matrices are denoted by  $C'_k$ 's. As proved in Appendix VII C, the overlap between the wave functions  $\Psi_e = \Phi_1 \wedge \cdots \wedge \Phi_n$  and  $\Psi'_e = \Phi'_1 \wedge \cdots \wedge \Phi'_n$  is given by the formula:

$$\langle \Psi_e | \Psi'_e \rangle = \langle \Phi_1 \wedge \cdots \wedge \Phi_n | \Phi'_1 \wedge \cdots \wedge \Phi'_n \rangle = \sum_{\substack{0 \leq N_{n,0}, \dots, N_{n,n} \leq n \\ \sum_{k=0}^n N_{n,k} = \sum_{k=0}^n k N_{n,k} = n}} (-1)^{N_{n,0}} \sum_{\sigma, \sigma' \in \mathfrak{S}_n} \prod_{k=1}^n \frac{\prod_{j=1}^{N_{n,k}} T_{N_{n,0}, \dots, N_{n,k-1}}^{N_{n,k}}(j, \sigma, \sigma')}{k^{N_{n,k}} N_{n,k}!} \quad (14)$$

with:

$$T_{N_{n,0}, \dots, N_{n,k-1}}^{N_{n,k}}(j, \sigma, \sigma') = \text{tr} \left[ C^\dagger_{\sigma \left( \sum_{p=0}^{k-1} p N_{n,p} + (j-1)k + 1 \right)} C'_{\sigma' \left( \sum_{p=0}^{k-1} p N_{n,p} + (j-1)k + 1 \right)} \cdots C^\dagger_{\sigma \left( \sum_{p=0}^{k-1} p N_{n,p} + jk \right)} C'_{\sigma' \left( \sum_{p=0}^{k-1} p N_{n,p} + jk \right)} \right], \quad (15)$$

if  $N_{n,k} \neq 0$ , and  $T_{N_{n,0}, \dots, N_{n,k-1}}^{N_{n,k}}(j, \sigma, \sigma') = 1$  by convention if  $N_{n,k} = 0$ .

The external sum is over the "partitions of the integer  $n$ ", i.e. the decompositions of  $n$  into a sum of (possibly repeated) natural integers. Each such partition is entirely determined by a set of integers  $(N_{n,k})_{k=1, \dots, n}$ , which correspond to the numbers of times integer  $k$  is found in the decomposition. That is to say, we can write:

$$n = \sum_{k=1}^n k N_{n,k} = \sum_{k=0}^n k N_{n,k} \quad . \quad (16)$$

The number  $N_{n,0}$ , which is introduced in the second equality, extends the definition to

$k = 0$  and depends linearly on the others  $N_{n,k}$ 's:

$$N_{n,0} = n - \sum_{k=1}^n N_{n,k} \quad . \quad (17)$$

It plays a particularly important rôle in the formula, as it determines the sign factor of each term.

The internal sum on permutations gives *a priori*  $(n!)^2$  terms. However, there are redundancies due to the invariance of a trace by circular permutation of its matrices, and by permutations of the traces of the same form, themselves. So, for a given partition of  $n$ , the number of *a priori* distinct products of traces to be calculated, reduces to

$$\frac{(n!)^2}{\prod_{k=1}^n k^{N_{n,k}} N_{n,k}!}.$$

These products of traces are decomposed into double products, as follows:

The most external product runs over an index  $k$  which determines the number of matrices ( $k$  matrices  $C_i$  and  $k$  matrices  $C'_j$ ) to be multiplied within each of the  $N_{n,k}$  trace factors (provided  $N_{n,k} \neq 0$ ) of the most internal product on the  $j$ -index.

**Example ( $n = 3$ ) :** We compute  $\langle \Phi_1 \wedge \Phi_2 \wedge \Phi_3 | \Phi'_1 \wedge \Phi'_2 \wedge \Phi'_3 \rangle = A_1 + A_2 + A_3$  , as follow:

$$\begin{aligned} \blacktriangleright 3=0+0+3: N_{3,0} = 2, N_{3,1} = 0, N_{3,2} = 0, N_{3,3} = 1 &\rightsquigarrow A_1 = \sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{\text{tr} \left[ C_{\sigma(1)}^\dagger C'_{\sigma'(1)} C_{\sigma(2)}^\dagger C'_{\sigma'(2)} C_{\sigma(3)}^\dagger C'_{\sigma'(3)} \right]}{3}. \\ \blacktriangleright 3=0+1+2: N_{3,0} = 1, N_{3,1} = 1, N_{3,2} = 1, N_{3,3} = 0 &\rightsquigarrow A_2 = - \sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{\text{tr} \left[ C_{\sigma(1)}^\dagger C'_{\sigma'(1)} \right] \text{tr} \left[ C_{\sigma(2)}^\dagger C'_{\sigma'(2)} C_{\sigma(3)}^\dagger C'_{\sigma'(3)} \right]}{2}. \\ \blacktriangleright 3=1+1+1: N_{3,0} = 0, N_{3,1} = 3, N_{3,2} = 0, N_{3,3} = 0 &\rightsquigarrow A_3 = \sum_{\sigma, \sigma' \in \mathfrak{S}_3} \frac{\text{tr} \left[ C_{\sigma(1)}^\dagger C'_{\sigma'(1)} \right] \text{tr} \left[ C_{\sigma(2)}^\dagger C'_{\sigma'(2)} \right] \text{tr} \left[ C_{\sigma(3)}^\dagger C'_{\sigma'(3)} \right]}{6}. \end{aligned}$$

Formula Eqs.(14),(15) can easily be checked in the particular case of an AGP against Eq.(8) of Ref.<sup>60</sup>. In this case, all geminal matrices are equal to one and the same matrix, say  $C_0$ :  $\forall k, C_k = C'_k = C_0$ . So, the sum over permutations gives just a factor  $(n!)^2$ . Moreover, by choosing the  $\phi_i$ 's to be the natural orbitals of the geminal function, we

can make this matrix diagonal, without loss of generality:  $\forall i, j, (C_0)_{i,j} = \delta_{i,j}\eta_i$ . So, all traces reduce to traces of powers of  $C_0$ , and  $\text{tr}[C_0^p] = \sum_{i=1}^m \eta_i^p$ . In the above example, we obtain

$$\langle \Phi_1 \wedge \Phi_2 \wedge \Phi_3 | \Phi_1 \wedge \Phi_2 \wedge \Phi_3 \rangle = 36 \times \left( \frac{\sum_{i=1}^m \eta_i^6}{3} - \frac{\sum_{i=1}^m \eta_i^2 \times \sum_{j=1}^m \eta_j^4}{2} + \frac{\sum_{i=1}^m \eta_i^2 \times \sum_{j=1}^m \eta_j^2 \times \sum_{k=1}^m \eta_k^2}{6} \right). \quad (18)$$

The coefficient of terms of the form  $\eta_i^6$  is found to be  $\frac{1}{3} - \frac{1}{2} + \frac{1}{6} = 0$ , and that of terms  $\eta_i^2 \eta_j^4$  with  $i \neq j$  is  $-\frac{1}{2} + 3 \times \frac{1}{6} = 0$ , so we are left with terms of the form  $\eta_i^2 \eta_j^2 \eta_k^2$  with  $i, j, k$  all distinct, arising from the last term of Eq.(18) only, with a combinatorial factor  $3!$ . Hence the final result:

$$\langle \Phi_1 \wedge \Phi_2 \wedge \Phi_3 | \Phi_1 \wedge \Phi_2 \wedge \Phi_3 \rangle = 36 \times \left( \sum_{1 \leq i < j < k \leq m} \eta_i^2 \eta_j^2 \eta_k^2 \right), \quad (19)$$

which is the expected result, since in Ref.<sup>60</sup>, both bra and ket AGP's are normalized by a factor  $n!$ , thus eliminating the factor  $36 = (3!)^2$ . A similar path can be followed to retrieve the formula for an arbitrary  $n$ .

At the other extreme, we can check Eqs.(14),(15) for the APSG case. The 1-orthogonality constraint imposed to the geminals amounts to setting

$$\forall i \neq j \quad C_i^\dagger C_j' = 0. \quad (20)$$

So, only one partition of  $n$  in the external sum gives a non-zero contribution, namely  $n = \underbrace{1 + \dots + 1}_{n \text{ times}}$ . Moreover, this constraint also imposes  $\sigma = \sigma'$  i.e. transforms the double sum on permutations into a single sum. The latter just gives a factor  $N_{n,1}! = n!$  that cancels out with the redundancy factor in the denominator. Therefore, we obtain the expected formula:

$$\langle \Phi_1 \wedge \dots \wedge \Phi_n | \Phi_1' \wedge \dots \wedge \Phi_n' \rangle = \text{tr}[C_1^\dagger C_1'] \text{tr}[C_2^\dagger C_2'] \dots \text{tr}[C_n^\dagger C_n']. \quad (21)$$

One can also retrieve the more general formula recently obtained for APIG<sup>61</sup>, which encompasses the case of RG states first worked out by Sklyanin<sup>62</sup>. We were not aware

of the latter work, when we first derived<sup>63,64</sup> formula (14), although it has a similar structure to the formula obtained by this author. In the cases of APIG, all matrices are diagonal, so their product is also a diagonal matrix whose  $i^{th}$  diagonal element is just the product of the  $i^{th}$  diagonal elements of the matrices in the product. Clearly, all matrices commute, so instead of intertwining bra and ket matrices in the traces, Eq.(15), we can gather the bra matrices to the left, and the ket matrices to the right. So, our traces become exactly the sum over  $i$  appearing in Eq.(69) of Ref.<sup>61</sup>. In the latter equation there is also a sign factor and a numerical factor. After performing the product over the  $r$  elements of a partition in Eq.(71) of Ref.<sup>61</sup> (where  $M$  is  $n$  in our notation), the sign factor reduces to  $(-1)^{M-r}$ , which is exactly the same as our global sign factor  $(-1)^{N_{n,0}}$ . Regarding the numerical factor, after performing the double summation on the symmetric group and the product over the  $N_{n,k}$   $k$ -cycles ( $k$  in our notation is  $q$  in Ref.<sup>61</sup>), we will obtain  $N_{n,k}! \times (k!)^{2N_{n,k}}$  identical product of traces over  $k$ -cycles. The  $N_{n,k}!$  comes from the commutations of the traces and simplifies with the denominator, and for each of the  $N_{n,k}$   $k$ -cycles a factor  $k! \times k!$  is due to the commutation of the  $k$  bra matrices inside the trace, on one hand, and of the  $k$ -ket matrices on the other hand. Taking into account the  $k^{N_{n,k}}$  in our denominator (related to the invariance by circular permutation of each trace), we obtain for each  $k$ -cycle a factor  $k! \times (k-1)!$ , which is exactly the factor in Eq.(69) of Ref.<sup>61</sup>.

#### IV. COMPUTATIONAL COST REDUCTION BY ORTHOGONALITY CONSTRAINTS

The number of partitions of  $n$  giving the number of terms in the external sum of Eq.(18), grows asymptotically as  $\frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$  when  $n$  goes to infinity. So, even if one disregards the internal sum over non-redundant permutations, the cost of calculating overlaps in the general case is rapidly prohibitive. Hence, the necessity of introducing constraints such as those reviewed in Section II. Here, we investigate permutationally invariant

2-orthogonality constraints already mentioned in conclusion of Ref.<sup>7</sup>.

### A. permutationally invariant 2-orthogonality constraints

We have seen in the previous section how choosing all geminals parallel (actually equal) to one another or all 1-orthogonal simplifies the overlap formula. Our aim is to propose less drastic inter-geminal orthogonality constraint, intermediate between the AGP (parallel) and the APSG (1-orthogonality) constraints.

Let us consider wave functions which are products of singlet or triplet geminals (i.e. with symmetric or antisymmetric associated matrices). We want to impose to the geminals  $\Phi_k$ 's what we call "permutationally invariant 2-orthogonality constraints":

$$\forall i, j, k \in \{1, \dots, n\} \text{ distinct, } \begin{cases} \langle \Phi_i | \Phi_j \rangle = 0 \\ \Phi_k \leftarrow (\Phi_i \wedge \Phi_j) = 0 \end{cases} \quad (22)$$

(the interior product by  $\Phi_k$ , noted " $\Phi_k \leftarrow$ ", plays in first quantization the same rôle as annihilation in second quantization) or in terms of matrices:

$$\forall i, j, k \in \{1, \dots, n\} \text{ distinct, } \begin{cases} \text{tr}(C_i^\dagger C_j) = 0 \\ C_i C_k^\dagger C_j + C_j C_k^\dagger C_i = 0 \end{cases} \quad (23)$$

Clearly Eq.(20) implies Eq.(23) so 1-orthogonality is at least as strong as permutationally invariant 2-orthogonality. However, the question arises whether there actually exist geminals which are permutationally invariant 2-orthogonal without being 1-orthogonal i.e. whether 1-orthogonality is strictly stronger than permutationally invariant 2-orthogonality.

For  $m = 1$ , there can only be one non-zero  $C_i$  satisfying the 2-orthogonal condition, so 1-orthogonality is also satisfied. To answer the question positively, let us consider the case  $m = 2$ , so that the  $C_i$ 's are  $2 \times 2$ -matrices. In this limit case,



$\forall i, j, k$  distinct,  $\Phi_i \wedge \Phi_j \neq 0 \implies \Phi_i \wedge \Phi_j \propto \varphi_1 \wedge \varphi_2 \wedge \bar{\varphi}_1 \wedge \bar{\varphi}_2 \implies \Phi_k = 0$ . So, to have at least three distinct matrices, we necessarily have  $\forall i \neq j, \Phi_i \wedge \Phi_j = 0$ . A general study<sup>65</sup> shows that one can have at most three singlet solutions of the form:

$$\begin{aligned} C_1 &= \begin{pmatrix} e^{-i\tau} & 0 \\ 0 & e^{i\tau} \end{pmatrix} z_1 \quad ; \quad C_2 = \begin{pmatrix} \rho e^{-i\tau} & 1 \\ 1 & -\rho e^{i\tau} \end{pmatrix} z_2 \quad ; \\ C_3 &= \begin{pmatrix} -e^{-i\tau} & \rho \\ \rho & e^{i\tau} \end{pmatrix} z_3 \quad , \quad \text{with} \quad \begin{cases} \rho \in \mathbb{R}_+ \\ \tau \in [0, \pi[ \end{cases} \quad , \end{aligned} \quad (24)$$

and  $z_1, z_2, z_3 \in \mathbb{C}^*$  unimportant factors. To this set one can add at most one triplet geminal matrix. A convenient choice corresponding to  $\rho = \tau = 0, z_1 = z_2 = 1, z_3 = -1$  gives the set:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad i\sigma_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad , \quad (25)$$

where  $\sigma_x, \sigma_y$  and  $\sigma_z$  are the well-known Pauli matrices. We will call the set  $\{I_2, \sigma_x, i\sigma_y, \sigma_z\}$  the "4-type set". One can trivially verify that the associated geminals are not 1-orthogonal.

Another non-maximal but useful set of solutions to the permutationally invariant 2-orthogonality conditions in dimension 2 is  $\{G_\theta, \sigma_x, i\sigma_y\}$ , where:

$$G_\theta = \begin{pmatrix} \sqrt{2} \sin \theta & 0 \\ 0 & \sqrt{2} \cos \theta \end{pmatrix} \quad , \quad \text{with} \quad \theta \in [0, \pi[ \quad (26)$$

whose associated geminal is:

$$\Phi_{G_\theta} = \sqrt{2} (\sin \theta \varphi_1 \wedge \bar{\varphi}_1 + \cos \theta \varphi_2 \wedge \bar{\varphi}_2) \quad \text{with} \quad \|\Phi_{G_\theta}\| = \sqrt{2} \quad . \quad (27)$$

Note that by playing with  $\theta$ , one can retrieve special matrices of interest:

$$\left\{ \begin{array}{l} G_0 = \sqrt{2} E_{22} \\ G_{\frac{\pi}{4}} = I_2 \\ G_{\frac{\pi}{2}} = \sqrt{2} E_{11} \\ G_{\frac{3\pi}{4}} = \sigma_z \end{array} \right. , \quad (28)$$

where  $E_{ij} = (\delta_{ki}\delta_{lj})_{kl}$ . It is important to also consider such "3-type sets" in order to break the degeneracy of the one-electron reduced density matrix (1RDM) eigenvalues. Note that all matrix types are either diagonal or anti-diagonal, so we partition them into two sets:  $\mathcal{D} = \{I_2, \sigma_z, G_\theta\}$  and  $\bar{\mathcal{D}} = \{\sigma_x, i\sigma_y\}$ .

One could study in the same fashion general solutions for increasing  $m$ -values. However, we prefer in the present article to limit ourselves to  $m$ -dimensional matrices built as block-diagonal matrices, each block being either a 1-dimensional (1D) block (1-orthogonal part corresponding to a geminal  $\varphi_i \wedge \bar{\varphi}_i$ ) or a 2-dimensional (2D) block constituted of a  $2 \times 2$ -matrix (whose elements correspond to geminal basis functions  $\varphi_i \wedge \bar{\varphi}_i, \varphi_i \wedge \bar{\varphi}_{i+1}, \varphi_{i+1} \wedge \bar{\varphi}_i, \varphi_{i+1} \wedge \bar{\varphi}_{i+1}$ ) belonging to the 4-type set or to a 3-type set for some value of  $\theta$  to be optimized.



Note that the permutationally invariant 2-orthogonality conditions further impose that

$$\forall i, j, k \in \{1, \dots, m\} \text{ distinct, } \begin{cases} \sum_{\substack{l \text{ such that} \\ B_i^l = B_j^l}} \bar{\lambda}_i^l \lambda_j^l + \sum_{\substack{l \text{ such that} \\ B_{\theta_l, i}^l = B_{\theta_l, j}^l}} \bar{\lambda}_{\theta_l, i}^l \lambda_{\theta_l, j}^l = 0 \\ \lambda_i^l \bar{\lambda}_k^l \lambda_j^l = 0 \text{ as soon as } \exists \{i', j'\} \subset \{i, j, k\}, B_{i'}^l = B_{j'}^l \end{cases} \quad (30)$$

For practical reason, we choose to rather enforce the following sufficient set of conditions, called "extended permutationally invariant 2-orthogonality" (EPI2O) conditions:

$$\forall i, j \in \{1, \dots, m\} \text{ distinct, } \forall l \in \{1, \dots, m'\}, \begin{cases} B_i^l = B_j^l \implies \lambda_i^l \lambda_j^l = 0 \\ B_{\theta_l, i}^l = B_{\theta_l, j}^l \implies \lambda_{\theta_l, i}^l \lambda_{\theta_l, j}^l = 0 \end{cases}, \quad (31)$$

In other words, two distinct geminals are not allowed to have the same matrix type at a given block position, corresponding to superscript, say  $l$ . So, in a 4-type 2D-block, there can be at most 4 non-zero  $\lambda_i^l$  coefficient, each in factor of one of the 4 different block types, and similarly, in a 3-type 2D-block, there can be at most 3 non-zero  $\lambda_{\theta_l, i}^l$  coefficient, each in factor of one of the 3 possible block types.

*Remark 1:* A pair of successive 1D-blocks associated with say  $\varphi_i \wedge \bar{\varphi}_i, \varphi_{i+1} \wedge \bar{\varphi}_{i+1}$  and occupied in a given geminal matrix, say  $C_k$ , can always be replaced by a 3-type 2D-block with a non-zero coefficient  $\lambda$  for this block, only in factor a  $G_\theta$  in  $C_k$ . So instead of optimizing two coefficients  $\lambda_i$  and  $\lambda_{i+1}$  in factor of the 1D-blocks, one optimizes  $\lambda$  and  $\theta$ , and identifies  $\lambda_i = \lambda \times \sqrt{2} \sin \theta, \lambda_{i+1} = \lambda \times \sqrt{2} \cos \theta$ .

## B. Expression of the overlap for the EPI2O 2D-block geminal ansatz

Let  $\Phi_1, \dots, \Phi_n$  and  $\Phi'_1, \dots, \Phi'_n$  be geminals verifying the EPI2O conditions with for all  $l$ , each  $C'_k$  in the ket having the same matrix form as their  $C_k$  counterpart in the bra (same block dimensions, types and zero coefficients at the same places). One can anticipate a simplification of the overlap formula Eq.(14) since in every block there will be at most traces of length 4 (product of 4 pairs of block-submatrix) for the cases of

spin-unrestricted 4-type- $2D$ -blocks, of length 3 for the cases of spin-restricted 4-type- $2D$ -blocks or spin-unrestricted 3-type- $2D$ -blocks, of length 2 for the cases of spin-restricted 3-type- $2D$ -blocks, or of length 1 for the cases of  $1D$ -blocks.

We shall prove that Eq.(14) becomes:

$$\langle \Phi_1 \wedge \cdots \wedge \Phi_n | \Phi'_1 \wedge \cdots \wedge \Phi'_n \rangle = \sum_{\substack{0 \leq j_1, \dots, j_n \leq m' \\ \text{distinct if non-zero}}} \zeta_{\Phi_1, \Phi'_1}^{j_1} \cdots \zeta_{\Phi_n, \Phi'_n}^{j_n} \quad (32)$$

$$\text{with } \zeta_{\Phi_u, \Phi'_u}^j = \begin{cases} \sum_{t=h_{u-1}+1}^{h_u} \overline{\lambda}_u^{-t} \lambda'_u{}^{-t} & \text{if } j = 0 \\ 2 \overline{\lambda}_{\theta_j, u}^j \lambda'_{\theta_j, u}{}^j & \text{if } j \in \{1, \dots, m''\} \\ 2 \overline{\lambda}_u^j \lambda'_u{}^j & \text{if } j \in \{m'' + 1, \dots, m'\} \end{cases} \quad (33)$$

**Direct demonstration:**

To prove the formula, we do not need to distinguish between the 3-type- $2D$ -blocks, and the 4-type- $2D$ -blocks, so we will omit the  $\theta$  label of the coefficients and block matrices in the 3-type cases. We decompose each geminal  $\Phi_i$  into its 1-orthogonal part and its EPI2O part:

$$\Phi_i = \Phi_i^0 + \sum_{x=1}^{m'} \Phi_i^x \quad (34)$$

where

$$\Phi_i^0 := \sum_{t=h_{i-1}+1}^{h_i} \lambda_i^{-t} \varphi_t \wedge \overline{\varphi}_t \quad , \quad (35)$$

and

$$\Phi_i^x = \sum_{p, q \in \{1, 2\}} \lambda_i^x (B_i^x)_{p, q} \varphi_{\mathfrak{J}(x, p)} \wedge \overline{\varphi}_{\mathfrak{J}(x, q)} \quad . \quad (36)$$

$\mathfrak{J}$  denotes the function that maps component 1 or 2 of block  $x \in \{1, \dots, m'\}$  to the corresponding orbital index:

$$\begin{cases} \mathfrak{J}(x, 1) = h_n + 2x - 1 \\ \mathfrak{J}(x, 2) = h_n + 2x \end{cases} \quad . \quad (37)$$

Then,

$$\bigwedge_{1 \leq i \leq n} \Phi_i = \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \sum_{1 \leq x_1, \dots, x_k \leq m'} \Phi_{i_1}^0 \wedge \dots \wedge \Phi_{i_{n-k}}^0 \wedge \Phi_{i_1}^{x_1} \wedge \dots \wedge \Phi_{i_k}^{x_k} \quad . \quad (38)$$

where  $\{i_1, \dots, i_{n-k}\} \cup \{\bar{i}_1, \dots, \bar{i}_k\} = \{1, \dots, n\}$ , and  $\bar{i}_1 < \dots < \bar{i}_k$ . If the same  $2D$ -block is chosen, say  $x_a = x_b$ , for two different geminals  $a$  and  $b$ , then the EPI2O conditions imply  $B_{i_a}^{x_a} \neq B_{i_b}^{x_b}$ , so that  $\Phi_{i_a}^{x_a} \wedge \Phi_{i_b}^{x_b} = 0$ . Therefore, the internal sum can be restricted to  $x_1, \dots, x_k$  all distinct.

The 1-orthogonality between distinct blocks, whether  $1D$  or  $2D$ -blocks, implies that each term in the sum must be paired with its exact counterpart in  $\bigwedge_{1 \leq i \leq n} \Phi'_i$  to give a non-zero contribution to the overlap. For the  $1D$ -part, since there is an exact correspondence between geminal indices and non-zero  $1D$ -blocks, 1-orthogonality is enough to associate a single partner for each block. In contrast, for a given  $2D$ -block, say  $x$ , there could be up to 4 ket-geminal indices compatible with say  $\Phi_{i_l}^x$ . However, only the index corresponding to the same index  $\bar{i}_l$  will give a non-zero contribution, since the EPI2O conditions imply that different indices correspond to different matrix types and by 2-orthogonality of the submatrices:  $\bar{i}_l \neq \bar{i}_{l'} \implies \langle \Phi_{i_l}^x | \Phi_{i_{l'}}^x \rangle = 0$ . So we are left with:

$$\langle \bigwedge_{1 \leq i \leq n} \Phi_i | \bigwedge_{1 \leq i \leq n} \Phi'_i \rangle = \sum_{k=1}^n \sum_{1 \leq i_1 < \dots < i_{n-k} \leq n} \sum_{\substack{1 \leq x_1, \dots, x_k \leq m' \\ \text{distinct}}} \prod_{j=1}^{n-k} \langle \Phi_{i_j}^0 | \Phi_{i_j}^0 \rangle \prod_{l=1}^k \langle \Phi_{i_l}^{x_l} | \Phi_{i_l}^{x_l} \rangle \quad .$$

Now, if we do not distinguish the different terms according to their number  $k$  of  $2D$ -blocks, we obtain formulas (32) and (33) where the index  $j$  refers to a  $1D$ -part, if  $j = 0$ , or to  $2D$ -block of index  $j$ , if  $j > 0$ .

In practice, it is often interesting to have  $1D$ -parts, to build a Hartree-Fock component in the geminal product wave function, for example, or to have less terms in the overlap formula. However, let us take advantage of Remark 1 to simplify the discussion and assume that we have only  $2D$ -blocks. Note that, if the number of orbital basis functions

is odd, one can always add an extra one at infinity. Then in Eq.(32), there will be only non-zero  $j$ -indices. Since the latter are necessarily all distinct, Eq.(32) can be recast in the form:

$$\begin{aligned}
\langle \Phi_1 \wedge \cdots \wedge \Phi_n | \Phi'_1 \wedge \cdots \wedge \Phi'_n \rangle &= \sum_{1 \leq j_1 < \cdots < j_n \leq m'} \sum_{\sigma \in \mathfrak{S}_n} \zeta_{\Phi_1, \Phi'_1}^{\sigma(j_1)} \cdots \zeta_{\Phi_n, \Phi'_n}^{\sigma(j_n)} \\
&= \sum_{1 \leq j_1 < \cdots < j_n \leq m'} \sum_{\sigma \in \mathfrak{S}_n} \zeta_{\Phi_{\sigma(1)}, \Phi'_{\sigma(1)}}^{j_1} \cdots \zeta_{\Phi_{\sigma(n)}, \Phi'_{\sigma(n)}}^{j_n} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \sum_{1 \leq j_1 < \cdots < j_n \leq m'} \zeta_{\Phi_{\sigma(1)}, \Phi'_{\sigma(1)}}^{j_1} \cdots \zeta_{\Phi_{\sigma(n)}, \Phi'_{\sigma(n)}}^{j_n}, \quad (39)
\end{aligned}$$

where  $\sigma$  runs over all permutations of  $n$ -indices,  $\mathfrak{S}_n$ . For a block  $y$ , let  $\mathcal{U}(y) = \{u \in \{1, \dots, n\}, \lambda_u^x \neq 0\}$  be the set of indices of the geminals which contain a submatrix with a non-zero coefficient for this block. In fact, in Eq.(39) there are always a large number of terms that are zero, since the permutation must fulfill  $\forall i \in \{1, \dots, n\}, \sigma(i) \in \mathcal{U}(j_i)$  otherwise  $\zeta_{\Phi_{\sigma(i)}, \Phi'_{\sigma(i)}}^{j_i} = 0$  and we recall that  $Card(\mathcal{U}(j_i)) \leq 4$ .

Let us set  $\mathcal{I}_n^{m'}(\sigma) = \sum_{1 \leq x_1 < \cdots < x_n \leq m'} \zeta_{\Phi_{\sigma(1)}, \Phi'_{\sigma(1)}}^{x_1} \cdots \zeta_{\Phi_{\sigma(n)}, \Phi'_{\sigma(n)}}^{x_n}$  and more generally for  $p$  indices  $u_1 < \cdots < u_p$ , and  $q$  block indices  $y_1 < \cdots < y_q$ , let us denote  $\mathcal{I}_{n \setminus \{u_1, \dots, u_p\}}^{m' \setminus \{y_1, \dots, y_q\}}(\sigma)$  the quantity:

$$\sum_{\substack{1 \leq x_1 < \cdots < x_{n-p} \leq m' \\ x_1, \dots, x_{n-p} \notin \{y_1, \dots, y_q\}}} \zeta_{\Phi_{\sigma(1)}, \Phi'_{\sigma(1)}}^{x_1} \cdots \zeta_{\Phi_{\sigma(u_1-1)}, \Phi'_{\sigma(u_1-1)}}^{x_{u_1-1}} \zeta_{\Phi_{\sigma(u_1+1)}, \Phi'_{\sigma(u_1+1)}}^{x_{u_1}} \cdots \zeta_{\Phi_{\sigma(u_p-1)}, \Phi'_{\sigma(u_p-1)}}^{x_{u_p-p}} \zeta_{\Phi_{\sigma(u_p+1)}, \Phi'_{\sigma(u_p+1)}}^{x_{u_p+1-p}} \cdots \zeta_{\Phi_{\sigma(n)}, \Phi'_{\sigma(n)}}^{x_{n-p}}$$

where  $q$  geminals and  $p$  blocks are excluded with respect to  $\mathcal{I}_n^{m'}(\sigma)$ . This allows us to write the following recursion formula:

$$\mathcal{I}_n^{m'}(\sigma) = \mathcal{I}_n^{m' \setminus \{y\}}(\sigma) + \sum_{\substack{k=1 \\ \sigma(k) \in \mathcal{U}(y)}}^n \zeta_{\Phi_{\sigma(k)}, \Phi'_{\sigma(k)}}^y \times \mathcal{I}_{n \setminus \{k\}}^{m' \setminus \{y\}}(\sigma), \quad (40)$$

which can be useful to compute the  $\mathcal{I}_n^{m'}(\sigma)$ 's efficiently.

### C. Expression of the 2RDM for the EPI2O 2D-block geminal ansatz

To optimize variationally the expectation value of the electronic Coulomb Hamiltonian, we will make use of its contracted form

$${}^2\hat{h} = \frac{1}{n-1} \left( T_1 + T_2 + \sum_i Z_i \times \left( \frac{1}{\|R_i - r_1\|} + \frac{1}{\|R_i - r_2\|} \right) \right) + \frac{1}{\|r_1 - r_2\|}, \quad (41)$$

where  $T_1$ ,  $T_2$  and  $r_1$ ,  $r_2$  are, respectively, the kinetic and position operators of electrons 1 and 2,  $Z_i$  the charge of nucleus  $i$ , and  $R_i$  the position operator of nucleus  $i$ . The expectation value over wave function  $\Psi_e$  is then computed according to:

$$E = \text{tr}({}^2\Gamma(\Psi_e) {}^2\hat{h}). \quad (42)$$

So, only the 2-electron reduced density matrix (2RDM),  ${}^2\Gamma(\Psi_e)$  is needed *in fine*.

In order to provide general expressions for the more general, 2-electron, reduced, transition matrix (2RTM) between wave functions  $\Psi_e$  and  $\Psi'_e$ , that we will denote  ${}^2\Gamma(\Psi_e, \Psi'_e)$  (the 2RDM expression will be obtained by suppressing the primes), let us generalize our notation to 1D-blocks (using negative integers to number them):

$$\forall i \in \{1, \dots, n\}, \quad \forall t \in \{h_{i-1} + 1, \dots, h_i\}, \quad \text{we set } \Phi_i^{-t} = \lambda_i^{-t} \varphi_t \wedge \bar{\varphi}_t \text{ and } \mathcal{U}(-t) = \{i\}.$$

For any block  $j$ , we note  $\bar{\mathcal{U}}(j)$  the complement of  $\mathcal{U}(j)$  in  $\{1, \dots, n\}$ . We also set for  $p$  geminal indices,  $1 \leq u_1 < \dots < u_p \leq n$ , and  $q$  2D-block indices

$$0 < y_1 < \dots < y_q < m',$$

$$\mathcal{S}_{n \setminus \{u_1, \dots, u_p\}}^{m' \setminus \{y_1, \dots, y_q\}} = \sum_{\substack{0 \leq x_1, \dots, x_{n-p} \leq m' \\ \text{all distinct if non-zero} \\ x_1, \dots, x_{n-p} \notin \{y_1, \dots, y_q\}}} \zeta_{\Phi_1, \Phi'_1}^{x_1} \cdots \zeta_{\Phi_{u_1-1}, \Phi'_{u_1-1}}^{x_{u_1-1}} \zeta_{\Phi_{u_1+1}, \Phi'_{u_1+1}}^{x_{u_1}} \cdots \zeta_{\Phi_{u_p-1}, \Phi'_{u_p-1}}^{x_{u_p-p}} \zeta_{\Phi_{u_p+1}, \Phi'_{u_p+1}}^{x_{u_p+1-p}} \cdots \zeta_{\Phi_n, \Phi'_n}^{x_{n-p}}. \quad (43)$$

Since  $\Psi_e$  and  $\Psi'_e$  are product of  $(S_z = 0)$ -geminals, the 2RTM can be decomposed into 4  $\hat{S}_z$ -adapted blocks that we denote simply (omitting the  $(\Psi_e, \Psi'_e)$  dependency)  ${}^2\Gamma^{\gamma, \delta}$  with



$\gamma, \delta \in \{\alpha, \beta\}$  referring to spin operator  $\hat{S}_z$  eigenstates.

Let us first consider  ${}^2\Gamma^{\alpha,\beta}$ , the  ${}^2\Gamma^{\beta,\alpha}$  block being obtained by symmetry. Denoting  $\mathbf{c}_i$  (respectively  $\bar{\mathbf{c}}_i$ ) the annihilation operator of spin-orbital  $\varphi_i$  (respectively  $\bar{\varphi}_i$ )

$$\begin{aligned} & \forall i_1, j_1, i_2, j_2 \in \{1, \dots, m\}, \\ {}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} &= \langle \bar{\mathbf{c}}_{j_1} \mathbf{c}_{i_1} \Psi_e | \bar{\mathbf{c}}_{j_2} \mathbf{c}_{i_2} \Psi'_e \rangle \Leftrightarrow \\ {}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} &= \langle \varphi_{i_1} \wedge \bar{\varphi}_{j_1} \lrcorner \Phi_1 \wedge \dots \wedge \Phi_n | \varphi_{i_2} \wedge \bar{\varphi}_{j_2} \lrcorner \Phi'_1 \wedge \dots \wedge \Phi'_n \rangle \\ &= \sum_{\substack{-h_n \leq x_1, \dots, x_n \leq m' \\ \text{all distinct} \\ -h_n \leq y_1, \dots, y_n \leq m' \\ \text{all distinct}}} \langle \varphi_{i_1} \wedge \bar{\varphi}_{j_1} \lrcorner \Phi_1^{x_1} \wedge \dots \wedge \Phi_n^{x_n} | \varphi_{i_2} \wedge \bar{\varphi}_{j_2} \lrcorner \Phi_1^{y_1} \wedge \dots \wedge \Phi_n^{y_n} \rangle. \end{aligned}$$

Note that the summation can be restricted to block  $x_i$  and  $y_i$  such that  $i \in \mathcal{U}(x_i) \cap \mathcal{U}(y_i)$  for all  $i \in \{1, \dots, n\}$ . We define the "block function",  $\mathfrak{B}$ , that associates to an orbital its block number:

$$\begin{aligned} \mathfrak{B} : \quad \{1, \dots, m\} &\longrightarrow \mathbb{Z}^* \\ i &\longmapsto \mathfrak{B}(i) = \begin{cases} -i & \text{if } i \leq h_n \\ \lfloor \frac{i-h_n+1}{2} \rfloor & \text{if } i > h_n \end{cases}. \end{aligned}$$

where  $\lfloor z \rfloor$  is the floor of  $z$ , and the function  $\mathfrak{C}$ , that associates to an orbital of a  $2D$ -block its position in the block:

$$\begin{aligned} \mathfrak{C} : \quad \{h_n + 1, \dots, m\} &\longrightarrow \{1, 2\} \\ i &\longmapsto \mathfrak{C}(i) = i - h_n - 2(\mathfrak{B}(i) - 1). \end{aligned}$$

Finally, we introduce the notation  $\bar{j}$  for the complementary of index  $j \in \{1, 2\}$ , that is to say,  $\bar{j} = 2$  if  $j = 1$ , and  $\bar{j} = 1$  if  $j = 2$ . Then for orbital  $i$  and geminal  $k$ , we associate the index  $\mathfrak{D}(i, k)$  defined as follow:

$$\begin{cases} \mathfrak{D}(i, k) = \mathfrak{C}(i) & \text{if } B_k^{\mathfrak{B}(i)} \in \mathcal{D}(\mathfrak{B}(i)) \\ \mathfrak{D}(i, k) = \overline{\mathfrak{C}(i)} & \text{if } B_k^{\mathfrak{B}(i)} \in \overline{\mathcal{D}}(\mathfrak{B}(i)) \end{cases}.$$

**Case 1:  $\mathfrak{B}(i_1) \neq \mathfrak{B}(j_1)$**

Two distinct blocks are broken in the bra, then, to have at least one ket with non-zero overlap, it is necessary that  $\mathfrak{B}(i_2) = \mathfrak{B}(i_1)$  and  $\mathfrak{B}(j_2) = \mathfrak{B}(j_1)$ . Moreover, these two blocks have to appear within both  $x_1, \dots, x_n$  and  $y_1, \dots, y_n$ . The remaining  $n - 2$  blocks in each set must be two by two identical because of 1-orthogonality between the blocks, and each pair must be associated with the same geminal because of the EPI2O conditions. This means also that the 2 broken blocks must be associated with the same pair of geminals in the bra and in the ket, of indices say  $(k_1, k_2)$ .

**Subcase 1a:  $\mathfrak{B}(i_1) < 0$  and  $\mathfrak{B}(j_1) < 0$**

Then, necessarily  $i_2 = i_1$ ,  $j_2 = j_1$  and  $\mathcal{U}(\mathfrak{B}(i_1)) \neq \mathcal{U}(\mathfrak{B}(j_1))$  otherwise  ${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} = 0$ .  $k_1$  is the unique element of  $\mathcal{U}(\mathfrak{B}(i_1))$  and  $k_2$  is the unique element of  $\mathcal{U}(\mathfrak{B}(j_1))$

$${}^2\Gamma_{i_1 j_1, i_1 j_1}^{\alpha, \beta} = \bar{\lambda}_{k_1}^{-i_1} \lambda'_{k_1}{}^{-i_1} \bar{\lambda}_{k_2}^{-j_1} \lambda'_{k_2}{}^{-j_1} \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m'}. \quad (44)$$

**Subcase 1b:  $\mathfrak{B}(i_1) < 0$  and  $\mathfrak{B}(j_1) > 0$**

Then  ${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} \neq 0$ , implies  $i_2 = i_1$  and  $k_1$  is the unique element of  $\mathcal{U}(\mathfrak{B}(i_1))$ . The block  $\mathfrak{B}(j_1)$  is associated with the same geminal  $k_2 \in \mathcal{U}(\mathfrak{B}(j_1))$  in the bra and in the ket. The block matrix  $B_{k_2}^{\mathfrak{B}(j_1)}$  being either diagonal or anti-diagonal, after annihilation of the  $\beta$ -spin-orbitals, one can only match the lone ket and bra  $\alpha$ -spin-orbitals if  $j_2 = j_1$ . Then:

$${}^2\Gamma_{i_1 j_1, i_1 j_1}^{\alpha, \beta} = \sum_{k_2 \in \mathcal{U}(\mathfrak{B}(j_1))} \bar{\lambda}_{k_1}^{-i_1} \lambda'_{k_1}{}^{-i_1} \bar{\lambda}_{k_2}^{\mathfrak{B}(j_1)} \lambda'_{k_2}{}^{\mathfrak{B}(j_1)} (B_{k_2}^{\mathfrak{B}(j_1)})_{\mathfrak{D}(j_1, k_2), \mathfrak{C}(j_1)}^2 \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(j_1)\}}$$

which can be decomposed as:

$$\begin{aligned} {}^2\Gamma_{i_1 j_1, i_1 j_1}^{\alpha, \beta} &= \sum_{k_2 \in \mathcal{D}(\mathfrak{B}(j_1))} \bar{\lambda}_{k_1}^{-i_1} \lambda'_{k_1}{}^{-i_1} \bar{\lambda}_{k_2}^{\mathfrak{B}(j_1)} \lambda'_{k_2}{}^{\mathfrak{B}(j_1)} (B_{k_2}^{\mathfrak{B}(j_1)})_{\mathfrak{C}(j_1), \mathfrak{C}(j_1)}^2 \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(j_1)\}} \\ &+ \sum_{k_2 \in \bar{\mathcal{D}}(\mathfrak{B}(j_1))} \bar{\lambda}_{k_1}^{-i_1} \lambda'_{k_1}{}^{-i_1} \bar{\lambda}_{k_2}^{\mathfrak{B}(j_1)} \lambda'_{k_2}{}^{\mathfrak{B}(j_1)} \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(j_1)\}}. \end{aligned} \quad (45)$$

where  $\mathcal{D}(j)$  (resp.  $\bar{\mathcal{D}}(j)$ ) is the subset of  $\mathcal{U}(j)$  collecting only the indices corresponding to geminals having a diagonal (resp. anti-diagonal) submatrix for block  $j$ .  $\forall j \in \{1, \dots, m'\}$ ,  $\mathcal{U}(j) = \mathcal{D}(j) \cup \bar{\mathcal{D}}(j)$ . Note that the factor  $(B_{k_2}^{\mathfrak{B}(j_1)})_{\mathfrak{C}(j_1), \mathfrak{C}(j_1)}^2$  on the first line of Eq.(45) will be 1 in the 4-type case, and either  $2\cos^2\theta$  or  $2\sin^2\theta$  according to  $\mathfrak{C}(j_1)$  being 1 or 2 in the 3-type case. On the second line, we used the fact that  $\forall k_2 \in \bar{\mathcal{D}}(\mathfrak{B}(j_1))$ ,  $(B_{k_2}^{\mathfrak{B}(j_1)})_{1,2}^2 = (B_{k_2}^{\mathfrak{B}(j_1)})_{2,1}^2 = 1$ .

Subcase 1c:  $\mathfrak{B}(i_1) > 0$  and  $\mathfrak{B}(j_1) < 0$

By symmetry with the previous subcase:

$$\begin{aligned} {}^2\Gamma_{i_1 j_1, i_1 j_1}^{\alpha, \beta} &= \sum_{k_1 \in \mathcal{D}(\mathfrak{B}(i_1))} \bar{\lambda}_{k_1}^{\mathfrak{B}(i_1)} \lambda'_{k_1}{}^{\mathfrak{B}(i_1)} (B_{k_1}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(i_1)}^2 \bar{\lambda}_{k_2}^{-j_1} \lambda'_{k_2}{}^{-j_1} \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(i_1)\}} \\ &+ \sum_{k_1 \in \bar{\mathcal{D}}(\mathfrak{B}(i_1))} \bar{\lambda}_{k_1}^{\mathfrak{B}(i_1)} \lambda'_{k_1}{}^{\mathfrak{B}(i_1)} \bar{\lambda}_{k_2}^{-j_1} \lambda'_{k_2}{}^{-j_1} \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(i_1)\}}. \end{aligned} \quad (46)$$

Subcase 1d:  $\mathfrak{B}(i_1) > 0$  and  $\mathfrak{B}(j_1) > 0$

This last case is more complicated than the previous ones, since there are 3 sub-subcases to consider:

- (i)  $(k_1, k_2) \in \mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(j_1)) \times \mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(j_1))$ ,
- (ii)  $(k_1, k_2) \in \mathcal{U}(\mathfrak{B}(i_1)) \times \bar{\mathcal{U}}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(j_1))$ ,
- (iii)  $(k_1, k_2) \in \mathcal{U}(\mathfrak{B}(j_1)) \times \mathcal{U}(\mathfrak{B}(i_1)) \cap \bar{\mathcal{U}}(\mathfrak{B}(j_1))$ .

Using the notation already defined, the following general formula can be obtained for case (i):

$$\begin{aligned}
{}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} &= \sum_{\substack{k_1, k_2 \in \mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(j_1)) \\ k_1 \neq k_2}} \mathcal{S}_{n \setminus \{k_1, k_2\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(j_1)\}} \left( \delta_{i_1, i_2} \delta_{j_1, j_2} \left( \bar{\lambda}_{k_1}^{\mathfrak{B}(i_1)} \lambda'_{k_1}{}^{\mathfrak{B}(i_1)} (B_{k_1}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{D}(i_1, k_1)}^2 \bar{\lambda}_{k_2}^{\mathfrak{B}(j_1)} \lambda'_{k_2}{}^{\mathfrak{B}(j_1)} \right. \right. \\
&\times \left. \left. (B_{k_2}^{\mathfrak{B}(j_1)})_{\mathfrak{D}(j_1, k_2), \mathfrak{C}(j_1)}^2 + \bar{\lambda}_{k_2}^{\mathfrak{B}(i_1)} \lambda'_{k_2}{}^{\mathfrak{B}(i_1)} (B_{k_2}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{D}(i_1, k_2)}^2 \bar{\lambda}_{k_1}^{\mathfrak{B}(j_1)} \lambda'_{k_1}{}^{\mathfrak{B}(j_1)} (B_{k_1}^{\mathfrak{B}(j_1)})_{\mathfrak{D}(j_1, k_1), \mathfrak{C}(j_1)}^2 \right) \right. \\
&+ \delta_{\mathfrak{D}(i_1, k_1), \mathfrak{D}(i_2, k_2)} \delta_{\mathfrak{D}(j_1, k_2), \mathfrak{D}(j_2, k_1)} \left( \bar{\lambda}_{k_1}^{\mathfrak{B}(i_1)} \lambda'_{k_2}{}^{\mathfrak{B}(i_1)} (B_{k_1}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{D}(i_1, k_1)} (B_{k_2}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_2), \mathfrak{D}(i_2, k_2)} \right. \\
&\times \left. \bar{\lambda}_{k_2}^{\mathfrak{B}(j_1)} \lambda'_{k_1}{}^{\mathfrak{B}(j_1)} (B_{k_2}^{\mathfrak{B}(j_1)})_{\mathfrak{D}(j_1, k_2), \mathfrak{C}(j_1)} B_{k_1}^{\mathfrak{B}(j_1)} \right)_{\mathfrak{D}(j_2, k_1), \mathfrak{C}(j_2)} \\
&+ \delta_{\mathfrak{D}(i_1, k_2), \mathfrak{D}(i_2, k_1)} \delta_{\mathfrak{D}(j_1, k_1), \mathfrak{D}(j_2, k_2)} \left( \bar{\lambda}_{k_2}^{\mathfrak{B}(i_1)} \lambda'_{k_1}{}^{\mathfrak{B}(i_1)} (B_{k_2}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{D}(i_1, k_2)} (B_{k_1}^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_2), \mathfrak{D}(i_2, k_1)} \right. \\
&\times \left. \bar{\lambda}_{k_1}^{\mathfrak{B}(j_1)} \lambda'_{k_2}{}^{\mathfrak{B}(j_1)} (B_{k_1}^{\mathfrak{B}(j_1)})_{\mathfrak{D}(j_1, k_1), \mathfrak{C}(j_1)} B_{k_2}^{\mathfrak{B}(j_1)} \right)_{\mathfrak{D}(j_2, k_2), \mathfrak{C}(j_2)} \Big). \tag{47}
\end{aligned}$$

The factor  $\delta_{i_1, i_2} \delta_{j_1, j_2}$  comes from the argument used before, when the same  $2D$ -block matrices are matched in the bra and in the ket. The other Kröneckers symbols can be worked out explicitly, according to  $i_1$  and  $i_2$  (resp.  $j_1$  and  $j_2$ ) being equal or not, and by detailing whether the  $2D$ -block matrices are of  $\mathcal{D}$  or  $\bar{\mathcal{D}}$ -type, as was done in Eqs.(45) and (46). This is not done for Eq.(47) to alleviate the presentation.

The formula for case (ii) can be deduced from Eq.(47) by eliminating the terms where either  $\bar{\lambda}_{k_1}^{\mathfrak{B}(j_1)}$  or  $\lambda'_{k_1}{}^{\mathfrak{B}(j_1)}$  appears, which leaves only the first term. Similarly, for case (iii), one must eliminate the terms where either  $\bar{\lambda}_{k_1}^{\mathfrak{B}(i_1)}$  or  $\lambda'_{k_1}{}^{\mathfrak{B}(i_1)}$  appears, which leaves only the second term on the second line. Care must be taken to count only once, the pairs  $\{k_1, k_2\}$  such that one element is in  $\bar{\mathcal{U}}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(j_1))$  and the other in  $\mathcal{U}(\mathfrak{B}(i_1)) \cap \bar{\mathcal{U}}(\mathfrak{B}(j_1))$ , which appear both in case (ii) and in case (iii).

### Case 2: $\mathfrak{B}(i_1) = \mathfrak{B}(j_1)$

No broken block in the bra, then, there cannot be broken block in the ket too, so necessarily  $\mathfrak{B}(i_2) = \mathfrak{B}(j_2)$ .

#### Subcase 2a: $\mathfrak{B}(i_1) \neq \mathfrak{B}(i_2)$

$\mathfrak{B}(i_1)$  has to appear within  $x_1, \dots, x_n$  but not  $\mathfrak{B}(i_2)$  since it will disappear from the ket

after annihilation of  $\varphi_{i_2} \wedge \bar{\varphi}_{j_2}$  in it. Symmetrically,  $\mathfrak{B}(i_2)$  has to appear within  $y_1, \dots, y_n$  but not  $\mathfrak{B}(i_1)$ . The remaining  $n - 1$  blocks in each set must be two by two identical because of 1-orthogonality between the blocks, and each pair must be associated with the same geminal because of the EPI2O conditions. This means also that  $\mathfrak{B}(i_1)$  in the bra and  $\mathfrak{B}(i_2)$  in the ket must be associated with the same geminal of index, say  $k$ . This is only possible if  $\mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(i_2)) \neq \emptyset$ . When this is the case, we obtain for  $2D$ -blocks:

$$\begin{aligned}
{}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} &= \delta_{i_1, j_1} \delta_{i_2, j_2} \sum_{k \in \mathcal{D}(\mathfrak{B}(i_1)) \cap \mathcal{D}(\mathfrak{B}(i_2))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(i_1)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(i_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(i_2)\}} \\
&+ (1 - \delta_{i_1, j_1}) \delta_{i_2, j_2} \sum_{k \in \bar{\mathcal{D}}(\mathfrak{B}(i_1)) \cap \mathcal{D}(\mathfrak{B}(i_2))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(j_1)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(i_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(i_2)\}} \\
&+ \delta_{i_1, j_1} (1 - \delta_{i_2, j_2}) \sum_{k \in \mathcal{D}(\mathfrak{B}(i_1)) \cap \bar{\mathcal{D}}(\mathfrak{B}(i_2))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(i_1)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(j_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(i_2)\}} \\
&+ (1 - \delta_{i_1, j_1}) (1 - \delta_{i_2, j_2}) \sum_{k \in \bar{\mathcal{D}}(\mathfrak{B}(i_1)) \cap \bar{\mathcal{D}}(\mathfrak{B}(i_2))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(j_1)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(j_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(i_2)\}}.
\end{aligned} \tag{48}$$

or more compactly (but at the price of summing terms that are zero),

$${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} = \sum_{k \in \mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(i_2))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(j_1)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(j_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1), \mathfrak{B}(i_2)\}}. \tag{49}$$

It is possible for some blocks to be  $1D$ . Then,  $\mathcal{U}(\mathfrak{B}(i_1)) \cap \mathcal{U}(\mathfrak{B}(i_2))$  can contain at most one geminal index, say  $k$ . If  $\mathfrak{B}(i_1) < 0$  and  $\mathfrak{B}(i_2) > 0$  then necessarily  $i_1 = j_1$  and :

$${}^2\Gamma_{i_1 i_1, i_2 j_2}^{\alpha, \beta} = \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_2)})_{\mathfrak{C}(i_2), \mathfrak{C}(j_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_2)\}}, \tag{50}$$

else if  $\mathfrak{B}(i_2) < 0$  and  $\mathfrak{B}(i_1) > 0$  then necessarily  $i_2 = j_2$  and:

$${}^2\Gamma_{i_1 j_1, i_2 i_2}^{\alpha, \beta} = \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(j_1)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1)\}}. \tag{51}$$

For both  $\mathfrak{B}(i_1) < 0$  and  $\mathfrak{B}(i_2) < 0$  to occur, it is necessary that  $h_k - h_{k-1} > 1$  as  $\mathfrak{B}(i_1), \mathfrak{B}(i_2) \in \{h_{k-1} + 1, \dots, h_k\}$ , then

$${}^2\Gamma_{i_1 i_1, i_2 i_2}^{\alpha, \beta} = \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_2)} \mathcal{S}_{n \setminus \{k\}}^{m'}.$$

Subcase 2b:  $\mathfrak{B}(i_1) = \mathfrak{B}(i_2)$

$\mathfrak{B}(i_1)$  has to appear within  $x_1, \dots, x_n$  and within  $y_1, \dots, y_n$ . The remaining  $n - 1$  blocks in each set must be two by two identical because of 1-orthogonality between the blocks, and each pair must be associated with the same geminal because of the EPI2O conditions. This means also that  $\mathfrak{B}(i_1)$  in the bra and in the ket must be associated with the same geminal of index, say  $k \in \mathcal{U}(\mathfrak{B}(i_1))$ . When this is the case, we obtain for  $2D$ -blocks:

$${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \beta} = \sum_{k \in \mathcal{U}(\mathfrak{B}(i_1))} \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_1)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_1), \mathfrak{C}(j_1)} (B_k^{\mathfrak{B}(i_1)})_{\mathfrak{C}(i_2), \mathfrak{C}(j_2)} \mathcal{S}_{n \setminus \{k\}}^{m' \setminus \{\mathfrak{B}(i_1)\}}.$$

If  $\mathfrak{B}(i_1) < 0$  then necessarily  $i_1 = j_1 = i_2 = j_2$  and let  $k$  be the unique element of  $\mathcal{U}(\mathfrak{B}(i_1))$ :

$${}^2\Gamma_{i_1 i_1, i_1 i_1}^{\alpha, \beta} = \bar{\lambda}_k^{\mathfrak{B}(i_1)} \lambda_k^{\mathfrak{B}(i_1)} \mathcal{S}_{n \setminus \{k\}}^{m'}.$$

Let us now turn to  ${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \alpha}$ , the case of  ${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\beta, \beta}$  being analogous. Here, the Pauli principle makes it sufficient to consider indices  $i_1, j_1, i_2, j_2 \in \{1, \dots, m\}$  such that  $i_1 < j_1$  and  $i_2 < j_2$ :

$${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\alpha, \alpha} = \sum_{\substack{-h_n \leq x_1, \dots, x_n \leq m' \\ \text{all distinct} \\ -h_n \leq y_1, \dots, y_n \leq m' \\ \text{all distinct}}} \langle \varphi_{i_1} \wedge \varphi_{j_1} \leftarrow \Phi_1^{x_1} \wedge \dots \wedge \Phi_n^{x_n} | \varphi_{i_2} \wedge \varphi_{j_2} \leftarrow \Phi_1^{y_1} \wedge \dots \wedge \Phi_n^{y_n} \rangle.$$

After annihilation of the  $\alpha$ -spin-orbitals, the lone  $\beta$ -spin-orbitals in the bra and in the ket must be paired, this imposes  $\mathfrak{B}(i_1) = \mathfrak{B}(i_2)$  and  $\mathfrak{B}(j_1) = \mathfrak{B}(j_2)$ .

### Case 1: $\mathfrak{B}(i_1) \neq \mathfrak{B}(j_1)$

Two distinct blocks are broken and must be associated with the same pair of geminals in the bra and in the ket, of indices say  $(k_1, k_2)$ . The two blocks being 1-orthogonal, the algebra is the same whether the spins of the orbitals in each block are the same or not. The formulas for  ${}^2\Gamma^{\alpha, \alpha}$  are the same as those for  ${}^2\Gamma^{\alpha, \beta}$ . However, only the cases where  $i_1 < j_1$  and  $i_2 < j_2$  are relevant.

### Case 2: $\mathfrak{B}(i_1) = \mathfrak{B}(j_1)$

It is not possible to annihilate two spin-orbitals of the same spin in a given block, since all the blocks only contains ( $S_z = 0$ )-geminals. So this case does not give any non-zero contribution.

The expressions obtained for  ${}^2\Gamma$  are extremely simple to compute, provided that we have an efficient way of calculating the  $\mathcal{S}_{n \setminus \{u_1, \dots, u_p\}}^{m' \setminus \{y_1, \dots, y_q\}}$ 's with  $p, q \leq 2$ .

### D. Expression of gradients for the EPI2O 2D-block geminal ansatz

In the EPI2O 2D-block geminal ansatz, there are three types of continuous parameters:

- (i) The coefficients defining the orbitals  $\phi_i$ 's in a given basis set. We will not attempt to optimize them in the present study.
- (ii) The block coefficients  $\lambda_k^x$ 's.
- (iii) In the case of 3-type-2D-blocks, the angles  $\theta_x$ 's.

In order to optimize the last two kinds, we provide formulas to compute the gradients with respect to their variations. We will assume that the  $\lambda_k^x$ 's are real, so that  $\bar{\lambda}_k^x = \lambda_k^x$ .

#### 1. Gradient with respect to block coefficients

We first note that all the formulas of the previous section for the  $2RTM$ , hence for the  $2RDM$ , are of the form:

$${}^2\Gamma_{i_1 j_1, i_2 j_2}^{\gamma_1, \gamma_2} = \sum_w \mathbf{B}_w \Lambda_w \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}, \quad (52)$$

where  $\mathbf{B}_w$  is a product of block matrix elements (reduced to 1 in the case of 1D-block), which can depend upon  $\theta_x$ 's but not upon  $\lambda_k^x$ 's;  $\Lambda_w$  is a product of block matrix coefficients,  $\lambda_k^x$ 's, which does not depend upon  $\theta_x$ 's;  $K_w$  is a set of at most two 2D-block indices and  $X_w$  is a set of at most two geminal indices. Like  $\Lambda_w$ , the quantity  $\mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}$

only depends upon the  $\lambda_k^x$ 's.

Taking the derivatives with respect to  $\lambda_k^x$  in Eq.(52) gives:

$$\frac{\partial^2 \Gamma_{i_1 j_1, i_2 j_2}^{\gamma_1, \gamma_2}}{\partial \lambda_k^x} = \sum_w \mathbf{B}_w \left( \frac{\partial \Lambda_w}{\partial \lambda_k^x} \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w} + \Lambda_w \frac{\partial \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}}{\partial \lambda_k^x} \right). \quad (53)$$

However, for a given  $w$ , at most one of the two terms can be non-zero:  $\frac{\partial \Lambda_w}{\partial \lambda_k^x}$  will be non-zero only if  $k \in K_w$ , whereas  $\frac{\partial \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}}{\partial \lambda_k^x}$  will be non-zero only if  $k \notin K_w$ .

Let us consider  $\frac{\partial \Lambda_w}{\partial \lambda_k^x}$ :

In "case 1", it will be non-zero only if  $k \in K_w$  and  $x \in \mathfrak{B}(i_1) \cup \mathfrak{B}(j_1)$ , then:

$$\frac{\partial \Lambda_w}{\partial \lambda_k^x} = \begin{cases} \frac{\Lambda_w}{\lambda_k^x} & \text{in the last two terms of case 1d(i)} \\ \frac{2\Lambda_w}{\lambda_k^x} & \text{in all other cases} \end{cases}. \quad (54)$$

In "case 2", it will be non-zero only if  $k \in K_w$  and  $x \in \mathfrak{B}(i_1) \cup \mathfrak{B}(i_2)$ , then:

$$\frac{\partial \Lambda_w}{\partial \lambda_k^x} = \begin{cases} \frac{\Lambda_w}{\lambda_k^x} & \text{if } \mathfrak{B}(i_1) \neq \mathfrak{B}(i_2) \\ \frac{2\Lambda_w}{\lambda_k^x} & \text{if } \mathfrak{B}(i_1) = \mathfrak{B}(i_2) \end{cases}. \quad (55)$$

Regarding  $\frac{\partial \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}}{\partial \lambda_k^x}$ , it will be non-zero only if  $k \notin K_w$ , then:

$$\frac{\partial \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}}{\partial \lambda_k^x} = \begin{cases} 2\lambda_k^x \mathcal{S}_{n \setminus K_w \cup \{k\}}^{m' \setminus X_w} & \text{if } x < 0 \\ 4\lambda_k^x \mathcal{S}_{n \setminus K_w \cup \{k\}}^{m' \setminus X_w} & \text{if } x > 0 \end{cases}. \quad (56)$$

Note that these gradients make sense only for coefficients  $\lambda_k^x$  such that  $k \in \mathcal{U}(x)$ .

## 2. Gradient with respect to angle parameters

In the case of a 3-type-2D-block of index  $x$ ,  $1 \leq x \leq m''$ , taking the derivatives with respect to  $\theta_x$  in Eq.(52) gives:

$$\frac{\partial^2 \Gamma_{i_1 j_1, i_2 j_2}^{\gamma_1, \gamma_2}}{\partial \theta_x} = \sum_w \frac{\partial \mathbf{B}_w}{\partial \theta_x} \Lambda_w \mathcal{S}_{n \setminus K_w}^{m' \setminus X_w}. \quad (57)$$



The quantity  $\frac{\partial \mathbf{B}_w}{\partial \theta_x}$  will be non-zero only if  $k \in K_w$ , and:

For case 1, if  $B_k^x \in \mathcal{D}(\mathfrak{B}(i_1)) \cup \mathcal{D}(\mathfrak{B}(j_1))$ , then

$$\frac{\partial \mathbf{B}_w}{\partial \theta_x} = \begin{cases} \text{in the last two terms of case 1d(i)} & \begin{cases} \cot \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{1,1} \text{ appears in } \mathbf{B}_w \\ -\tan \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{2,2} \text{ appears in } \mathbf{B}_w \end{cases} \\ \text{in all other cases} & \begin{cases} 2 \cot \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{1,1}^2 \text{ appears in } \mathbf{B}_w \\ -2 \tan \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{2,2}^2 \text{ appears in } \mathbf{B}_w \end{cases} \end{cases} \quad (58)$$

For case 2, if  $B_k^x \in \mathcal{D}(\mathfrak{B}(i_1)) \cup \mathcal{D}(\mathfrak{B}(i_2))$ , then,

$$\frac{\partial \mathbf{B}_w}{\partial \theta_x} = \begin{cases} \text{if } \mathfrak{B}(i_1) \neq \mathfrak{B}(i_2) & \begin{cases} \cot \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{1,1} \text{ appears in } \mathbf{B}_w \\ -\tan \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{2,2} \text{ appears in } \mathbf{B}_w \end{cases} \\ \text{if } \mathfrak{B}(i_1) = \mathfrak{B}(i_2) & \begin{cases} 2 \cot \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{1,1}^2 \text{ appears in } \mathbf{B}_w \\ -2 \tan \theta_x \mathbf{B}_w & \text{if } (B_k^x)_{2,2}^2 \text{ appears in } \mathbf{B}_w \\ (\cot \theta_x - \tan \theta_x) \mathbf{B}_w & \text{if } (B_k^x)_{1,1} (B_k^x)_{2,2} \text{ appears in } \mathbf{B}_w \end{cases} \end{cases} \quad (59)$$

## V. PROOF OF CONCEPT

We have implemented our new geminal ansatz in the Nice-branch of the TONTO<sup>66</sup> quantum chemistry code. We have already noted (see Remark 1) that two 1D-blocks with non-zero coefficients associated with the same geminal can always be replaced by a 3-type-2D-block with a  $G_\theta$ -submatrix attributed to that geminal, the optimization of the two 1D  $\lambda$ -coefficients amounting to that of the  $\lambda$ -coefficient and the  $\theta$ -angle of the  $G_\theta$ -submatrix. So, provided that the Hilbert subspaces associated with the geminals of an APSG wave function have all even dimensions, we can reproduce it exactly by using 3-type-2D-blocks and  $G_\theta$ -submatrices only. In the present study, we want to show within the simple frame of 3-type-2D-blocks, that the accuracy of the new ansatz goes beyond that of APSG. Although we have implemented the general ansatz in TONTO, we

choose to focus here on EPI2O-geminals made of 3-type-2D-blocks only. It is sufficient to encompass APSG in many cases, and to go beyond it, as we shall see.

Moreover, we will limit our study to the singlet spin-restricted case. So, for each block, we will have a choice to allocate non-zero coefficients between only two types: a  $G_\theta$ -type or a  $\sigma_x$ -type. Assuming  $m$  even, there will be  $\frac{m}{2}$ -blocks. For each of these blocks, the position of the  $G_\theta$ -type will be fixed to allow an APSG approximate solution to be reproduced according to our previous remark. Then, it will remain to fix the position of the  $\sigma_x$ -types. There are  $n-1$  positions left for each block, so, in total, there are  $(n-1)^{\frac{m}{2}}$  possibilities to allocate the  $\sigma_x$  among the  $n$  geminals.

The correlation energy gained with respect to the APSG model being due to  $\sigma_x$ -submatrices, it comes exclusively from non-zero seniority configurations. The difference between the full configuration interaction FCI energy of a molecular system and the lowest possible energy in the seniority-zero approximation, that is to say the doubly-occupied configuration interaction DOCI energy, being often maximum around the minimum of its Born-Oppenheimer potential energy surface (PES), (see for example Fig. (8a) of Ref.<sup>2</sup>), we will concentrate our study on molecular systems at their ground state equilibrium geometry.

To study the dependency upon the number of geminals with a fixed number of orbitals, we have chosen to study diatomics of the second row of the periodic tables. A minimal basis set being too small to display significant dynamic correlation effects, we have used double zeta with polarization basis, resulting in 28 spherical orbitals, that is to say 14 2D-blocks, for each system. Actually, we also report results for a triple zeta basis, with 30 2D-blocks to investigate basis set effects. However, the flexibility offered by such a large number of blocks was too large to be explored in this preliminary study. So, we just transposed the strategy for the allocation of block types that was derived from our

double zeta examples. We have chosen to limit ourselves to homonuclear systems, so that respecting their symmetry helps also to restrict the possibilities of allocating group types to the different geminals. The sample of selected molecules although small covers single, double and triple bond systems.

The results are presented in Tab. I. The difference between our best spin-restricted 3-type-2D-block, so actually 2-type-2D-block,  $E_{\{G_\theta, \sigma_x\}}^0$  energy, and the APSG one,  $E_{APSG}^0$ , is of course due to the release of the strong orthogonality constraint, but can also be seen as resulting from the release of the seniority-zero constraint. So, it provides a lower bound to the correlation energy due to seniority non-zero configurations. We find this difference significant since it is more than the hundredth of Hartree for all the systems displayed. We cannot claim that the  $E_{\{G_\theta, \sigma_x\}}^0$  energies are the lowest that can be obtained within our model, for we have not investigated all possible allocations of the  $\sigma_x$ -type non-zero coefficients among the geminals, neither have we tested all the permutations of the molecular orbitals to form the 2D blocks. However, these results are sufficient to establish a proof of concept that the EPI2O model can be significantly more accurate than the APSG one.

Let us explain the strategy we have followed to obtain the results of Tab. I. We start from APSG calculations. The optimized APSG geminals are numbered according to their associated energies in increasing order. So, geminals 1 and 2 are always degenerate core geminals, and the others with higher energies, numbered from 3 to 7 in the case of  $N_2$ , from 3 to 8 for  $O_2$ , and from 3 to 9 for  $F_2$ , are valence geminals. The geminals are very much localized and are further classified into 3 sets according to the partial charges associated with the two atoms. The group where, say atom A, has partial charge  $\approx 2$  and atom B  $\approx 0$ , is labelled "2/0", group "0/2" corresponds to the reverse, and group "1/1" to a partial charge of  $\approx 1$  on each atom. For  $N_2$ , geminals 1 and 4 belong to group 2/0 while their degenerate partners, geminals 2 and 5 to group 0/2, the three

geminals 3, 6 and 7 to group 1/1. A set of natural orbitals is associated with each 1-orthogonal geminal and we rely on their populations to pair them into  $2D$ -blocks. That is to say, when the number of natural orbitals attached to a geminal is even, say  $2k$ , we make  $k$   $2D$ -blocks ordered according to decreasing absolute values of geminal expansion coefficients. When the number is odd, we transfer the orbital having the expansion coefficient with the least absolute value, to another geminal of the same partial charge group when possible. We try also to only apply permutations that preserve geminal degeneracies. For example, the dimensions of the Hilbert subspaces of the geminals in the  $N_2$  cc-pVTZ case, are: 2 2 10 14 14 9 9. So, block 1 is made of orbitals 1 and 2 associated with a  $G_\theta$ -type for geminal 1, block 2 is made of orbitals 3 and 4 associated with a  $G_\theta$ -type for geminal 2, blocks 3 to 7 are made of orbitals 5 to 14 all associated with  $G_\theta$ -types for geminal 3. However, since degenerate geminals 6 and 7 have odd numbers of natural orbitals, we moved orbitals 51 and 60, which are their least occupied orbitals, to positions 15 and 16 respectively, to make a sixth  $2D$ -block associated with a  $G_\theta$ -type for geminal 3, the latter being in the same 1/1 group as geminals 6 and 7. Then, the next 7, (7, 4, 4 respectively)  $2D$ -blocks are associated with  $G_\theta$ -types for geminals 4, (5, 6, 7, respectively). The reordered geminal's natural orbitals together with the submatrix-type allocation for all the calculations presented in this article, can be found in input and output files available in the "tonto/tests/geminal" github repository. The allocation of the  $\sigma_x$ -types is done in the following way. For cc-pVDZ  $N_2$  and  $O_2$  cases, we found that only the  $2D$ -blocks associated with the natural orbitals with the highest occupancy in each geminal are improving the energy when a non-zero coefficient is attributed to a  $\sigma_x$ -type. This is why the number of parameters actually used has a value equal to  $m+n$  in Tab. I, and not the maximum of  $\frac{3m}{2}$  for the 2-type model. For example, for the  $N_2$  cc-pVDZ case, we used only  $28 + 7 = 35$  parameters out of 42. We find that the best energy is obtained by allocating the  $\sigma_x$ -types to geminals of the same group. In the  $N_2$  example: we allocate a  $\sigma_x$ -type to geminal 4 for the dominant block of geminal 1 and conversely, to geminal 5 for the dominant block of geminal 2 and conversely, to geminal

7 for the dominant blocks of geminals 3 and 6, and to geminal 6 for the dominant block of geminal 7. In contrast, for  $F_2$  one can gain 641 microHartrees by using all of the 42 available parameters. This seems system-dependent rather than basis-dependent, since for  $N_2$  in the cc-pVTZ basis, the energy obtained by using only  $\sigma_x$ -types for dominant blocks, that is to say 67 parameters, is just 9 microHartrees above the lowest energy we have managed to find for this system. The latter can be obtained with only 70 parameters out of 90. Usually, several different allocations are able to give the same energy and they may correspond to different numbers of terms in the overlap formula. The allocations with non-zero coefficients evenly spread among the geminals have more terms than those where the non-zero terms are packed in one or a few geminals. The "nb. terms in overlap" and CPU times displayed in Tab. I, correspond to allocations minimizing these two correlated quantities. Note that the CPU times provided in the last line of Tab. I are only indicative of the order of magnitude, as no statistical processing has been performed.

## VI. CONCLUSION

After an overview of existing APG-based methods, we have derived a general formula for the overlap between APG wave functions. It exhibits the full combinatorics complexity of APG-models and helps to understand the simplifications arising in known APG-constrained model. For example, the sum over the partitions of the number of geminals,  $n$ , reduces to a single term when the 1-orthogonality constraint is enforced: the term with only 1-cycle traces. In the model studied in the previous section, we have gone one step further: imposing the constraint that at most two coefficients per  $2D$ -block can be non-zero, implies that the sum over the partitions of  $n$  are limited to the  $(\lfloor \frac{n}{2} \rfloor + 1)$  terms having at most 2-cycle traces, where  $\lfloor k \rfloor$  denotes the integer part of  $k$ . More generally, in the case of spin-unrestricted 3-type  $2D$  blocks or spin-restricted 4-type  $2D$  blocks, there can be at most 3 non-zero coefficients per block, therefore the

partitions of  $n$  are limited to terms involving at most 3-cycle traces. In the same way, in the case of spin-unrestricted 4-type  $2D$  blocks the sum on partitions of  $n$  is limited to terms having at most 4-cycle traces.

The permutationally invariant 2-orthogonality constraints, proposed in Section IV A, restrict the double summation over permutations in the overlap formula. Considering again the spin-restricted 3-type  $2D$  blocks of the previous section, the 2-orthogonality constraint (upper relation in Eq.(23)) implies that  $\sigma = \sigma'$  on the geminal indices in 1-cycle traces in Eq.(14). The second relation allows one to restrict the 2-cycle traces to those having  $C$ -matrices indices in increasing order. More precisely, let  $\text{tr}\left[C_{\sigma(i)}^\dagger C_{\sigma'(i)} C_{\sigma(j)}^\dagger C_{\sigma'(j)}\right]$  be a general 2-cycle trace, its invariance by circular permutation permits to choose  $\sigma(i) < \sigma(j)$  without loss of generality, and, if  $\sigma'(i)$ ,  $\sigma(j)$  and  $\sigma'(j)$  are all distinct, the constraint  $C_{\sigma'(i)} C_{\sigma(j)}^\dagger C_{\sigma'(j)} = -C_{\sigma'(j)} C_{\sigma(j)}^\dagger C_{\sigma'(i)}$  permits to choose  $\sigma'(i) < \sigma'(j)$  without loss of generality, provided one keeps track of the sign factor. (If  $\sigma(i) \in \{\sigma'(i), \sigma'(j)\}$ , then the normalization of the geminals amounting to  $\forall i, C_i^\dagger C_i = Id$ , the 2-cycle trace actually reduces to a 1-cycle one.) So, the combination of the restriction on the number of non-zero coefficients per block, together with the permutationally invariant 2-orthogonality (PI2O) conditions, keeps under control the combinatorial growth of the number of terms appearing in the overlap formula. This constitutes the extended PI2O (EPI2O) 2D-block geminal model, denoted as EPI2O-APG.

We have obtained explicit formulas for the quantities necessary to implement the EPI2O-APG ansatz and provided a proof of principle that the latter model is able to give strictly lower energies than the APSG one. This has been accomplished in a simple version of the EPI2O-APG ansatz where the factors breaking the strong orthogonality constraint were all of seniority greater than 0. So, the extra correlation energy with respect to APSG was solely due to the removal of the seniority-zero constraint. The latter has

been enforced in all recently proposed geminal approaches. The present study appears to be the first affordable geminal method taking into account seniority non-zero contributions. Although these contributions do not dominate quantitatively the energy, there are usually not homogeneous across a PES. So, our model could improve such properties as dissociation energies, energy barriers or harmonic frequencies. This will be investigated in future works, together with an in-depth assessment of the computational performances of the EPI2O geminal ansatz.

The present work raises many questions and calls for many further developments. What are the best orbital block partition and the best submatrix-type allocation to geminals? The number of possibilities is huge. A number of empirical rules have been postulated in this study and partially confirmed. However, to deal with this flexibility, at least partially, it would still be desirable to optimize the molecular orbitals, so as to automatically create the best orbital block partition for a given submatrix-type allocation. In addition, according to the variational principle, this would necessarily improve the results. Based on the example of orbital optimization for AP1roG wave functions<sup>19,21</sup>, we can expect this improvement to be non negligible. The introduction of 4-type  $2D$ -blocks should also be investigated, as this would add non strongly orthogonal components of seniority zero in the wave function, not present in the simplified model we have applied in the present article. So, this would partially separate out non 1-orthogonal contributions from non-zero seniority ones. Another direction for future developments is the implementation of the ansatz on hybrid quantum-classical computers. The structure of the wave function in terms of Pauli matrices (or  $2 \times 2$ -diagonal matrices) hints to its encoding by qubits associated with pairs of spin-orbitals, as proposed in<sup>58,59</sup>. In contrast with these works, the division by two of the number of qubits necessary for the wave function encoding would not be possible due to the release of the seniority-zero constraint. However, there could be a quantum advantage in calculating the energy for a EPI2O-APG wave function within a VQE-type algorithm<sup>67</sup>. A quantum walk search

algorithm could also be considered to explore the different submatrix-type allocations.

## ACKNOWLEDGEMENTS

This work has been supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01, and the grant CARMA ANR-12-BS01-0017. Prof. V. Rassolov is acknowledged for fruitful discussions and his help in calculating 1-orthogonal starting guesses. We are also thankful to J.-M. Lacroix and R. Ruelle for their support in using the cluster of the Lab. Dieudonné

## VII. APPENDIX

### A. Generalized seniority

The concept relies on a given decomposition of the one-electron Hilbert space,  $\mathcal{H}$ , into a direct sum of Hilbert subspaces called “shells”,  $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \dots \oplus \mathcal{H}_n$ , (where the dimension of  $\mathcal{H}_i$  is  $\dim \mathcal{H}_i := d_i$ ) or more conveniently, in practice, on a partition of the spin-orbital basis set into subsets,  $(\chi_{i,j})_{j \in \{1, \dots, d_i\}}$ , spanning the shells of the decomposition,  $i \in \{1, \dots, n\}$  being the shell index<sup>68</sup>.

The generalized seniority number of a Slater determinant built over these spin-orbitals is simply the number of incomplete shells present in the determinant i.e. shells that are neither empty nor fully-occupied. More precisely, let  $\chi_{i_1, j_1} \wedge \dots \wedge \chi_{i_1, j_{k_1}} \wedge \dots \wedge \chi_{i_p, j_p} \wedge \dots \wedge \chi_{i_p, j_{k_p}}$  be such a Slater determinant (where  $\wedge$  the intrinsiquely antisymmetrical, “Grassmann” or “exterior” product<sup>69</sup>), one counts 1 for every  $k_m$  such that  $0 < k_m < d_m$  (assuming all spin-orbitals are distinct otherwise the determinant is zero). This defini-



tion extends over linear combinations of Slater determinants having same generalized seniority number. Then, note that the definition is invariant under arbitrary rotations (or more generally, unitary transformations) of the spin-orbitals within the same shell. Hence, it is legitimate to talk of seniority with respect to the shell decomposition and not just with respect to the spin-orbital partition.

When there are  $m$  orbitals, one can retrieve the usual notion of seniority<sup>70–73</sup> by decomposing the Hilbert space into  $m$  two-dimensional Hilbert subspaces,  $\mathcal{H} = \mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_m$ , where each  $\mathcal{H}_i$  is spanned by two spin-orbitals of spin- $\alpha$  and spin- $\beta$ , respectively. However, other partitions not only based on spin-degeneracy, can prove physically relevant, such as partitions into atomic shell orbitals (hence the name we have coined), or into spatially degenerate molecular orbitals.

## B. $p$ -orthogonality

Let  $p > 0$  be an integer,  $\Psi_1$  an  $n_1$ -electron wave function, and  $\Psi_2$  an  $n_2$ -electron one, with  $n_1 \geq p$  and  $n_2 \geq p$ . Let us consider, their  $p$ -particle reduced density matrices,  ${}^p\Gamma_1$  and  ${}^p\Gamma_2$  respectively. We call “ $p$ -external space of  $\Psi_i$ ” the  $p$ -particle Hilbert subspace spanned by the unoccupied  $p$ -particle natural wave functions, that is to say, by the eigenfunctions of  ${}^p\Gamma_i$  associated with the 0 eigenvalue, for  $i \in \{1, 2\}$ , and “ $p$ -internal space of  $\Psi_i$ ”, noted  $\mathcal{I}[\Psi_i]$  the orthogonal subspace. In other words,  $\mathcal{I}[\Psi_i]$  is the Hilbert space spanned by the partially occupied  $p$ -particle natural wave functions. We will say that  $\Psi_1$  and  $\Psi_2$  are  $p$ -orthogonal if and only if their  $p$ -internal spaces are orthogonal:  $\mathcal{I}[\Psi_1] \perp \mathcal{I}[\Psi_2]$ . Note that 1-orthogonality is nothing but the so-called “strong orthogonality”<sup>74–76</sup>, and when  $n_1 = n_2 = n$ ,  $n$ -orthogonality is just the usual concept of orthogonality. For any integers  $0 < p \leq q \leq n$ , we have the important property that  $p$ -orthogonality implies  $q$ -orthogonality, which means that the lesser the integer  $p$ , the

stronger the  $p$ -orthogonality. An earlier definition of this concept is due to S. Wilson<sup>77</sup>.

### C. Proof of the general APG overlap formula

We will prove our formula by mathematical induction on  $k \geq 1$ , assuming that the equality is true for a particular  $k$  (the case  $k = 1$  being trivial). For the step  $(k + 1)$ , we want to calculate the following scalar product:

$$\langle \Phi_1 \wedge \cdots \wedge \Phi_{k+1} | \Phi'_1 \wedge \cdots \wedge \Phi'_{k+1} \rangle \quad .$$

In order to use the induction hypothesis, we transform the last exterior product in the bra into a right co-product, by Hopf duality. The left-hand side becomes:

$$\Phi_1 \wedge \cdots \wedge \Phi_k \otimes \Phi_{k+1} \quad ,$$

while the action of the co-product on the right-hand side gives the following expression:

$$\begin{aligned} & \Phi'_1 \wedge \cdots \wedge \Phi'_k \otimes \Phi'_{k+1} + \sum_{u=1}^k \Phi'_1 \wedge \cdots \wedge \Phi'_{u-1} \wedge \Phi'_{k+1} \wedge \Phi'_{u+1} \wedge \cdots \wedge \Phi'_k \otimes \Phi'_u \\ & - \sum_{1 \leq \alpha, \beta, \gamma, \delta \leq m} \sum_{1 \leq u < v \leq k+1} [(C'_u)_{\gamma, \beta} (C'_v)_{\alpha, \delta} + (C'_v)_{\gamma, \beta} (C'_u)_{\alpha, \delta}] \times \\ & \quad \times \Phi'_1 \wedge \cdots \wedge \Phi'_{u-1} \wedge \Phi'_{u+1} \wedge \cdots \wedge \Phi'_{v-1} \wedge \Phi'_{v+1} \wedge \cdots \wedge \Phi'_{k+1} \wedge \varphi_\alpha \wedge \bar{\varphi}_\beta \otimes (\varphi_\gamma \wedge \bar{\varphi}_\delta) \quad . \end{aligned}$$

With this new formulation, we can develop each expression and apply the induction hypothesis. The scalar product can now be written as a sum of two distinct terms:  $K_1$ , composed by the terms in which the geminals of the right-hand side term have not been cut by the co-product, and  $K_2$ , formed by the other terms.

For  $K_1$ , we obtain:

$$\begin{aligned} K_1 &= \langle \Phi_1 \wedge \cdots \wedge \Phi_k | \Phi'_1 \wedge \cdots \wedge \Phi'_k \rangle \langle \Phi_{k+1} | \Phi'_{k+1} \rangle + \sum_{u=1}^k \langle \Phi_1 \wedge \cdots \wedge \Phi_k | \bigwedge_{\substack{1 \leq l \leq k+1 \\ l \neq u}} \Phi'_l \rangle \langle \Phi_{k+1} | \Phi'_u \rangle \\ &= \langle \Phi_1 \wedge \cdots \wedge \Phi_k | \Phi'_1 \wedge \cdots \wedge \Phi'_k \rangle \text{tr}(C_{k+1}^\dagger C'_{k+1}) + \sum_{u=1}^k \langle \Phi_1 \wedge \cdots \wedge \Phi_k | \bigwedge_{\substack{1 \leq l \leq k+1 \\ l \neq u}} \Phi'_l \rangle \text{tr}(C_{k+1}^\dagger C'_u) \quad . \end{aligned}$$

Applying the induction hypothesis on the two terms of  $K_1$  (for the second one, by

$$K_1 = \sum_{\substack{0 \leq N_{k,0}, \dots, N_{k,k} \leq k \\ \sum_{i=0}^k N_{k,i} = \sum_{i=0}^k i N_{k,i} = k}} Q_{N_{k,0}, \dots, N_{k,k}} \sum_{u=1}^k \text{tr}(C_{k+1}^\dagger C'_u) \sum_{\sigma, \sigma' \in \mathfrak{S}_k} \prod_{i=1}^k P_{N_{k,0}, \dots, N_{k,i-1}}^{N_{k,i}}(\sigma, \tau_{u, k+1} \sigma') \quad , \quad (60)$$

with  $\tau_{u, k+1}$  the transposition  $(u, k+1) \in \mathfrak{S}_{k+1}$ .

For  $K_2$  we have:

$$\begin{aligned} K_2 &= - \sum_{1 \leq \alpha, \beta, \gamma, \delta \leq m} \sum_{1 \leq u < v \leq k+1} [(C'_u)_{\gamma, \beta} (C'_v)_{\alpha, \delta} + (C'_v)_{\gamma, \beta} (C'_u)_{\alpha, \delta}] \langle \Phi_1 \wedge \dots \wedge \Phi_k | \left( \bigwedge_{\substack{1 \leq l \leq k+1 \\ l \neq u, v}} \Phi'_l \right) \wedge \varphi_\alpha \wedge \overline{\varphi_\beta} \rangle \langle \Phi_{k+1} | \varphi_\gamma \wedge \overline{\varphi_\delta} \rangle \\ &= - \sum_{1 \leq u < v \leq k+1} \sum_{1 \leq \alpha, \beta \leq m} \left[ \langle \Phi_1 \wedge \dots \wedge \Phi_k | \left( \bigwedge_{\substack{1 \leq l \leq k+1 \\ l \neq u, v}} \Phi'_l \right) \wedge \varphi_\alpha \wedge \overline{\varphi_\beta} \rangle \times \right. \\ &\quad \left. \times \sum_{1 \leq \gamma, \delta \leq m} \left( (C'_v)_{\alpha, \delta} \langle \Phi_{k+1} | \varphi_\gamma \wedge \overline{\varphi_\delta} \rangle (C'_u)_{\gamma, \beta} + (C'_u)_{\alpha, \delta} \langle \Phi_{k+1} | \varphi_\gamma \wedge \overline{\varphi_\delta} \rangle (C'_v)_{\gamma, \beta} \right) \right] \\ &= - \sum_{1 \leq u < v \leq k+1} \sum_{1 \leq \alpha, \beta \leq m} \langle \Phi_1 \wedge \dots \wedge \Phi_k | \left( \bigwedge_{\substack{1 \leq l \leq k+1 \\ l \neq u, v}} \Phi'_l \right) \wedge \varphi_\alpha \wedge \overline{\varphi_\beta} \rangle \left[ (C'_v C_{k+1}^\dagger C'_u)_{\alpha\beta} + (C'_u C_{k+1}^\dagger C'_v)_{\alpha\beta} \right] \quad . \end{aligned}$$

Let  $E_{\alpha\beta}$  be the matrix where only the coefficient of line  $\alpha$  and column  $\beta$  is non-zero and equal to 1. We can use the induction hypothesis by replacing  $\Phi'_u$  with  $\varphi_\alpha \wedge \overline{\varphi_\beta}$  and, for  $v \neq k+1$ ,  $\Phi'_v$  with  $\Phi'_{k+1}$  (i.e. by substituting  $C'_u$  with  $E_{\alpha\beta}$  and, for  $v \neq k+1$ ,  $C'_v$  with  $C'_{k+1}$ ). To take advantage of the simple form of  $E_{\alpha\beta}$ , we have to rewrite our recursion formula as:

$$\begin{aligned} &\sum_{\substack{0 \leq N_{k,0}, \dots, N_{k,k} \leq k \\ \sum_{i=0}^k N_{k,i} = \sum_{i=0}^k i N_{k,i} = k \\ N_{k, x_u} \geq 1}} Q_{N_{k,0}, \dots, N_{k,k}} \sum_{x_u=1}^k N_{k, x_u} x_u \sum_{\substack{\sigma, \sigma' \in \mathfrak{S}_k \\ \sigma'(L_{x_u}) = u}} \text{tr} \left[ \left( \prod_{l=L_{x_u-1} + (N_{k, x_u} - 1)x_u + 1}^{L_{x_u} - 1} C_{\sigma(l)}^\dagger C'_{\sigma'(l)} \right) C_{\sigma(L_{x_u})}^\dagger E_{\alpha\beta} \right] \times \\ &\quad \times P_{N_{k,0}, \dots, N_{k, x_u-1}}^{N_{k, x_u} - 1}(\sigma, \sigma') \prod_{\substack{1 \leq i \leq k \\ i \neq x_u}} P_{N_{k,0}, \dots, N_{k, i-1}}^{N_{k, i}}(\sigma, \sigma') \quad , \end{aligned}$$

with  $L_{x_u} = \sum_{p=0}^{x_u} p N_{k, p}$ .

Note that a multiplying factor has appeared. Indeed, since we have distinguished the trace containing  $E_{\alpha\beta}$ , our formula is no longer completely symmetric under permutations. We had to multiply by  $N_{k,x_u}$ , the number of possible positions of that specific trace of length  $2x_u$  in its product with the other traces of same length (consequence of the commutativity of multiplication in  $\mathbb{K}$ ), and then by  $x_u$ , which is the number of possible locations for  $E_{\alpha\beta}$  in our trace (consequence of the invariance of the trace under circular permutations of matrices).

We also notice that  $N_{k,x_u}$  cannot be zero since it is related to the trace containing  $E_{\alpha\beta}$ , which necessarily exists. This trace is actually reduced to a single matrix element:

$$\mathrm{tr} \left[ \left( \prod_{l=L_{x_u-1}+(N_{k,x_u}-1)x_u+1}^{L_{x_u}-1} C'_{\sigma(l)} C'_{\sigma'(l)} \right) C'_{\sigma(L_{x_u})} E_{\alpha\beta} \right] = \left[ \left( \prod_{l=L_{x_u-1}+(N_{k,x_u}-1)x_u+1}^{L_{x_u}-1} C'_{\sigma(l)} C'_{\sigma'(l)} \right) C'_{\sigma(L_{x_u})} \right]_{\beta\alpha} . \quad (61)$$

Then, a *bona fide* trace re-appears in  $K_2$  since sum:

$$\sum_{1 \leq \alpha, \beta \leq m} [(C'_v C'_{k+1} C'_u)_{\alpha\beta} + (C'_u C'_{k+1} C'_v)_{\alpha\beta}] \left[ \left( \prod_{l=L_{x_u-1}+(N_{k,x_u}-1)x_u+1}^{L_{x_u}-1} C'_{\sigma(l)} C'_{\sigma'(l)} \right) C'_{\sigma(L_{x_u})} \right]_{\beta\alpha}$$

is nothing but:

$$\mathrm{tr} \left[ C'_{\sigma(L_{x_u})} (C'_v C'_{k+1} C'_u + C'_u C'_{k+1} C'_v) \prod_{l=L_{x_u-1}+(N_{k,x_u}-1)x_u+1}^{L_{x_u}-1} C'_{\sigma(l)} C'_{\sigma'(l)} \right] .$$

By using the induction hypothesis,  $K_2$  can be written:

$$\begin{aligned} K_2 = & - \sum_{1 \leq u < v \leq k+1} \sum_{x_u=1}^k \sum_{\substack{0 \leq N_{k,0}, \dots, N_{k,k} \leq k \\ \sum_{i=0}^k N_{k,i} = \sum_{i=0}^k i N_{k,i} = k \\ N_{k,x_u} \geq 1}} Q_{N_{k,0}, \dots, N_{k,k}} N_{k,x_u} x_u \sum_{\substack{\sigma, \sigma' \in \mathfrak{S}_k \\ \sigma'(L_{x_u})=u}} P_{N_{k,0}, \dots, N_{k,x_u-1}}^{N_{k,x_u}-1} (\sigma, \tau_{v,k+1} \sigma') \times \\ & \times \mathrm{tr} \left[ C'_{\sigma(L_{x_u})} (C'_v C'_{k+1} C'_u + C'_u C'_{k+1} C'_v) \prod_{l=L_{x_u-1}+(N_{k,x_u}-1)x_u+1}^{L_{x_u}-1} C'_{\sigma(l)} C'_{\tau_{v,k+1} \sigma'(l)} \right] \prod_{\substack{1 \leq i \leq k \\ i \neq x_u}} P_{N_{k,0}, \dots, N_{k,i-1}}^{N_{k,i}} (\sigma, \tau_{v,k+1} \sigma') , \end{aligned} \quad (62)$$

with  $\tau_{v,k+1}$  the transposition  $(v, k+1) \in \mathfrak{S}_{k+1}$ .

Note that the trace which is factored in the general term  $K_2$  and whose length has increased from  $2x_u$  to  $2(x_u + 1)$  actually represents two traces of  $C$ -matrix products since we have kept the symmetrized form  $(C'_v C'_{k+1}{}^\dagger C'_u + C'_u C'_{k+1}{}^\dagger C'_v)$  inside in order to reduce the size of the  $K_2$ -expression.

$K_1$  and  $K_2$  have now a very similar form. However, we still have to replace the sum on partitions of  $k$  with one on partitions of  $k + 1$ . To perform this, we define several mappings.

First, let  $\mathcal{F}_{k,0}$  denote the following mapping:

$$\begin{aligned} \mathcal{F}_{k,0} : \quad \mathbb{N}^{k+1} &\longrightarrow \mathbb{N} \times \mathbb{N}^* \times \mathbb{N}^{k-1} \times \{0\} \\ (N_{k,0}, \dots, N_{k,k}) &\longmapsto (N_{k,0}, N_{k,1} + 1, N_{k,2}, \dots, N_{k,k}, 0) \quad . \end{aligned}$$

For  $1 \leq i \leq k - 1$ , we define  $\mathcal{F}_{k,i}$  as:

$$\begin{aligned} \mathcal{F}_{k,i} : \mathbb{N}^i \times \mathbb{N}^* \times \mathbb{N}^{k-i} &\longrightarrow \mathbb{N}^* \times \mathbb{N}^i \times \mathbb{N}^* \times \mathbb{N}^{k-i-1} \times \{0\} \\ (N_{k,0}, \dots, N_{k,k}) &\longmapsto (N_{k,0} + 1, N_{k,1}, \dots, N_{k,i-1}, N_{k,i} - 1, N_{k,i+1} + 1, N_{k,i+2}, \dots, N_{k,k}, 0) \quad . \end{aligned}$$

Finally, we introduce  $\mathcal{F}_{k,k}$  as follows:

$$\begin{aligned} \mathcal{F}_{k,k} : \quad \mathbb{N}^k \times \mathbb{N}^* &\longrightarrow \mathbb{N}^* \times \mathbb{N}^k \times \{1\} \\ (N_{k,0}, \dots, N_{k,k}) &\longmapsto (N_{k,0} + 1, N_{k,1}, \dots, N_{k,k-1}, N_{k,k} - 1, 1) \quad . \end{aligned}$$

These mappings transform a partition of  $k$  into a partition of  $k + 1$ . Moreover, they are invertible on their image and their inverses allow to map a partition of  $k + 1$  onto a partition of  $k$ :

$$\begin{aligned} \mathcal{F}_{k,0}^{-1} : \quad \mathbb{N} \times \mathbb{N}^* \times \mathbb{N}^{k-1} \times \{0\} &\longrightarrow \mathbb{N}^{k+1} \\ (N_{k+1,0}, \dots, N_{k+1,k}, 0) &\longmapsto (N_{k+1,0}, N_{k+1,1} - 1, N_{k+1,2}, \dots, N_{k+1,k}) \quad ; \end{aligned}$$

$$\mathcal{F}_{k,i}^{-1} : \mathbb{N}^* \times \mathbb{N}^i \times \mathbb{N}^* \times \mathbb{N}^{k-i-1} \times \{0\} \longrightarrow \mathbb{N}^i \times \mathbb{N}^* \times \mathbb{N}^{k-i} \quad (1 \leq i \leq k-1)$$

$$(N_{k+1,0}, \dots, N_{k+1,k}, 0) \longmapsto (N_{k+1,0} - 1, N_{k+1,1}, \dots, N_{k+1,i-1}, N_{k+1,i} + 1, N_{k+1,i+1} - 1, N_{k+1,i+2}, \dots, N_{k+1,k}) \quad ;$$

$$\mathcal{F}_{k,k}^{-1} : \quad \mathbb{N}^* \times \mathbb{N}^k \times \{1\} \longrightarrow \mathbb{N}^k \times \mathbb{N}^*$$

$$(N_{k+1,0}, \dots, N_{k+1,k}, 1) \longmapsto (N_{k+1,0} - 1, N_{k+1,1}, \dots, N_{k+1,k-1}, N_{k+1,k} + 1) \quad .$$

If we consider a partition of  $k$  which has  $q$  non-zero  $N_{k,i}$ 's ( $i \geq 1$ ), then we can apply exactly  $q + 1$  distinct mappings  $\mathcal{F}_{k,i}$  (one for each non-zero  $N_{k,i}$ , and  $\mathcal{F}_{k,0}$  is always applicable), each giving a different partition of  $k + 1$ .

Conversely, let  $(N_{k+1,0}, \dots, N_{k+1,k+1})$  be an arbitrary partition of  $k + 1$ . Consider  $q \geq 1$  and  $i_1 < \dots < i_q$  such as the  $N_{k+1,i_j+1}$ 's are the non-zero integers determining our partition of  $k + 1$  (we deduce the value  $N_{k+1,0}$ ). This partition can then stem from exactly  $q$  different partitions of  $k$ , denoted by  $(N_{k,0}^{i_j}, \dots, N_{k,k}^{i_j})$ :

$$\left\{ \begin{array}{l} (N_{k,0}^{i_1}, \dots, N_{k,k}^{i_1}) = \mathcal{F}_{k,i_1}^{-1}(N_{k+1,0}, \dots, N_{k+1,k+1}) \\ (N_{k,0}^{i_2}, \dots, N_{k,k}^{i_2}) = \mathcal{F}_{k,i_2}^{-1}(N_{k+1,0}, \dots, N_{k+1,k+1}) \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ (N_{k,0}^{i_q}, \dots, N_{k,k}^{i_q}) = \mathcal{F}_{k,i_q}^{-1}(N_{k+1,0}, \dots, N_{k+1,k+1}) \end{array} \right. .$$

So, we have three possible forms for the preimages of  $(N_{k+1,0}, \dots, N_{k+1,k+1})$ .

(i) If  $i_1 = 0$ , then:

$$(N_{k,0}^0, \dots, N_{k,k}^0) = (N_{k+1,0}, N_{k+1,1} - 1, N_{k+1,2}, \dots, N_{k+1,k}) \quad . \quad (63)$$

We see from Eq.(60), that such a pre-image will arise from  $K_1$ .

(ii) If  $i_q = k$ , then  $q = 1$  and:

$$(N_{k,0}^k, \dots, N_{k,k}^k) = (N_{k+1,0} - 1, N_{k+1,1}, \dots, N_{k+1,k-1}, N_{k+1,k} + 1) \quad . \quad (64)$$

(iii) Finally, for the general case  $1 \leq i_j \leq k - 1$ :

$$(N_{k,0}^{i_j}, \dots, N_{k,k}^{i_j}) = (N_{k+1,0}-1, N_{k+1,1}, \dots, N_{k+1,i_j-1}, N_{k+1,i_j}+1, N_{k+1,i_j+1}-1, N_{k+1,i_j+2}, \dots, N_{k+1,k}) \quad . \quad (65)$$

We see from Eq.(62), that these two last cases will arise from the terms in  $K_2$ .

So, all the products of traces required for our partition of  $k + 1$ ,  $(N_{k+1,0}, \dots, N_{k+1,k+1})$ , will appear either in  $K_1$  or in  $K_2$ ,  $C_{k+1}^\dagger$  being inserted in the trace that has been factored out from the product of traces, and the transpositions insuring that  $C'_{k+1}$  will be inserted in all possible factors of the remaining product of traces. Now, we just have to examine the coefficients coming from the different pre-images. Let us denote by  $Q_{i_j}$  the factor coming from the  $(N_{k,0}^{i_j}, \dots, N_{k,k}^{i_j})$  pre-image. The products of traces related to the latter are those, where  $C_{k+1}^\dagger$  appears in a trace of length  $2(i_j + 1)$ . We have to show that by transforming the summations in  $K_1$  or  $K_2$  into sums over  $\mathfrak{S}_{k+1}$ , these factors are all equal to the expected normalization factor:  $Q_{N_{k+1,0}, \dots, N_{k+1,k+1}}$ .

Let us first consider  $Q_0$  (only present if  $i_1 = 0$ ). The terms for which  $Q_0$  is factored come from  $K_1$ . To find the value of  $Q_0$  from  $Q_{N_{k,0}^0, \dots, N_{k,k}^0}$ , we have to examine how the summations over  $\sigma$ ,  $\sigma'$  and  $u$  can be compacted into a double summation over  $\mathfrak{S}_{k+1}$ . To each  $\sigma' \in \mathfrak{S}_k$ , corresponds  $k + 1$  distinct permutations  $\sigma'_{k+1,u} \in \mathfrak{S}_{k+1}$ :

$$\begin{cases} \sigma'_{k+1,u}(\sigma'^{-1}(u)) = k + 1 \\ \sigma'_{k+1,u}(k + 1) = u \\ \sigma'_{k+1,u}(l) = \sigma'(l) \text{ for all } l \in \{1, \dots, k\} \setminus \{\sigma'^{-1}(u)\} \end{cases} \quad . \quad (66)$$

In this way, the summation over  $u$  from 1 to  $k + 1$  and the one over  $\sigma' \in \mathfrak{S}_k$  can be condensed into a sum over  $\mathfrak{S}_{k+1}$  without changing the normalization factor already present. To extend the summation over  $\sigma$  from  $\mathfrak{S}_k$  to  $\mathfrak{S}_{k+1}$ , we have to divide by the number of possibilities to interleave the new trace of length 2 (the one containing  $C_{k+1}^\dagger$ )

within the  $(N_{k+1,1} - 1)$  other traces, (that is to say by  $(N_{k+1,1} - 1) + 1$ ). So, we obtain:

$$\begin{aligned}
Q_0 &= \underbrace{Q_{N_{k+1,0}, N_{k+1,1}-1, N_{k+1,2}, \dots, N_{k+1,k}}}_{\text{factor } Q_{N_{k,0}^0, \dots, N_{k,k}^0} \text{ appearing in } K_1} \times \frac{1}{\underbrace{N_{k+1,1}}_{\text{additional factor}}} \\
&= \frac{(-1)^{N_{k+1,0}}}{\prod_{i=1}^{k+1} i^{N_{k+1,i}} N_{k+1,i}!} \\
&= Q_{N_{k+1,0}, \dots, N_{k+1,k+1}} \quad .
\end{aligned}$$

We proceed likewise with the other  $Q_{i_j}$ 's, which will be factored in the terms coming from  $K_2$  when  $x_u = i_j$  and also when  $k = i_j$ . The  $i_j$ 's are in fact the only possible values for  $x_u$  since, at a fixed partition of  $k$ , the values of  $x_u$  giving non-zero terms in the sum are those corresponding to non-zero  $N_{k,x_u}$ . This time, it is the double summation over  $u$  and  $v$  in  $K_2$  which will allow us to transform the sum over  $\sigma' \in \mathfrak{S}_k$  into a sum over  $\mathfrak{S}_{k+1}$  without additional normalization factor. To each permutation  $\sigma' \in \mathfrak{S}_k$ , we can associate the following permutations  $\sigma'_{k+1,u,v} \in \mathfrak{S}_{k+1}$ :

$$\left\{ \begin{array}{l} \sigma'_{k+1,u,v}(\sigma'^{-1}(v)) = k + 1 \text{ if } v \neq k + 1 \\ \sigma'_{k+1,u,v}(L_{i_j}) = u \\ \sigma'_{k+1,u,v}(k + 1) = v \\ \sigma'_{k+1,u,v}(l) = \sigma'(l) \text{ for all } l \in \{1, \dots, k\} \setminus \{\sigma'^{-1}(v), L_{i_j}\} \end{array} \right. \quad . \quad (67)$$

These permutations are pairwise distinct when we consider all the possible values for  $u$  and  $v$  and all the permutations  $\sigma' \in \mathfrak{S}_k$ . In  $K_2$ , the double summation on  $(u, v)$  contains  $\frac{k(k+1)}{2}$  values and the sum over  $\sigma' \in \mathfrak{S}_k$  contains  $(k-1)!$  different permutations as  $\sigma'(L_{i_j}) = u$  is fixed. Therefore, we arrive at  $\frac{(k-1)!k(k+1)}{2} = \frac{(k+1)!}{2}$  distinct permutations in  $\mathfrak{S}_{k+1}$ . This number has to be multiplied by two, since, as already noticed, the trace of length  $2(i_j + 1)$  containing  $C_{k+1}^\dagger$  actually counts as two (the previous  $\frac{(k+1)!}{2}$  permutations can be composed with transposition  $\tau_{u,v}$  to account for the symmetrized  $C$ -matrix product). So, we have a total of  $(k+1)!$  distinct permutations, which means that we have all the permutations belonging to  $\mathfrak{S}_{k+1}$ . Regarding the extension of the summation over  $\sigma$



to  $\mathfrak{S}_{k+1}$ , we have to divide by the number of ways to interleave the new trace of length  $2(i_j + 1)$  (containing  $C_{k+1}^\dagger$ ) with the  $(N_{k+1,i_j+1} - 1)$  other traces, but also by a factor  $(i_j + 1)$ , which arises from the invariance of the trace containing  $C_{k+1}^\dagger$  of length  $2(i_j + 1)$  under circular permutation. We obtain as required:

$$\begin{aligned}
Q_{i_j} &= \underbrace{-Q_{N_{k+1,0}-1, N_{k+1,1}, \dots, N_{k+1,i_j-1}, N_{k+1,i_j}+1, N_{k+1,i_j+1}-1, N_{k+1,i_j+2}, \dots, N_{k+1,k}}}_{\text{factor } -Q_{N_{k,0}^{i_j}, \dots, N_{k,k}^{i_j}} \text{ appearing in } K_2} (N_{k+1,i_j} + 1) i_j \times \underbrace{\frac{1}{N_{k+1,i_j+1}(i_j + 1)}}_{\text{additional factor}} \\
&= \frac{(-1)^{N_{k+1,0}}}{\prod_{i=1}^{k+1} i^{N_{k+1,i}} N_{k+1,i}!} \\
&= Q_{N_{k+1,0}, \dots, N_{k+1,k+1}} \quad .
\end{aligned}$$

To resume our achievements: the sum over partitions of  $k$  have been transformed into a sum over partitions of  $k + 1$ , and all the other sums have been condensed into a double sum over  $\mathfrak{S}_{k+1}$  while the normalization factors turned into the expected  $Q_{N_{k+1,0}, \dots, N_{k+1,k+1}}$  factors. All the terms in  $K_1$  and  $K_2$  have been consumed along the way. This implies that  $\langle \Phi_1 \wedge \dots \wedge \Phi_{k+1} | \Phi'_1 \wedge \dots \wedge \Phi'_{k+1} \rangle = K_1 + K_2$  verifies our induction property for Step  $(k + 1)$ . This concludes our proof by induction of the general APG overlap formula.

## REFERENCES

- <sup>1</sup>Matthias Troyer, Uwe-Jens Wiese, Physical Review Letters **94**, 170201 (2005).
- <sup>2</sup>Charles-Émile Fecteau, Samuel Cloutier, Jean-David Moisset, Jérémy Boulay, Patrick Bultinck, Alexandre Faribault, Paul A. Johnson, The Journal of Chemical Physics **156**, 194103 (2022).
- <sup>3</sup>Rishab Dutta, Guo P. Chen, Thomas M. Henderson, and Gustavo E. Scuseria, The Journal of Chemical Physics, **154**, 114112 (2021).
- <sup>4</sup>Haochuan Wei, Eric Neuscamman, J. Chem. Phys. **149**, 184106 (2018).
- <sup>5</sup>Andrey M. Tokmachev, Int. J. Quantum Chem. **116**, 265-269 (2016).

- <sup>6</sup>Patrick Cassam-Chenaï, *Physical Review* **A77**, 032103 (2008).
- <sup>7</sup>Patrick Cassam-Chenaï, Vitaly Rassolov, *Chemical Physics Letters* **487**, 147-152 (2010).
- <sup>8</sup>R. McWeeny, B. Sutcliffe, *Proc. Roy. Soc. (London)* **A273**, 103-116 (1963).
- <sup>9</sup>Patrick Cassam-Chenaï, *The Journal of Chemical Physics* **124**, 194109 (2006).
- <sup>10</sup>Patrick Cassam-Chenaï, Giovanni Granucci, *Chemical Physics Letters* **450**, 151-155 (2007).
- <sup>11</sup>Paul A. Johnson, Peter A.Limacher, Taewon D. Kim, Michael Richer, Ramón Alain Miranda-Quintana, Farnaz Heidar Zadeh, Paul W. Ayers, Patrick Bultinck, Stijn De Baerdemacker, Dimitri Van Neck, *Computational and Theoretical Chemistry* **1116**, 207-219 (2017).
- <sup>12</sup>D. M. Silver, *J. Chem. Phys.* **50**, p5108 (1969).
- <sup>13</sup>P. A Limacher, *J. Chem. Phys.* **145**, 194102 (2016).
- <sup>14</sup>Paul A. Johnson, Paul W. Ayers, Peter A.Limacher, Stijn De Baerdemacker, Dimitri Van Neck, Patrick Bultinck, *Computational and Theoretical Chemistry* **1003**, 101-113 (2013).
- <sup>15</sup>Pawel Tecmer, Katharina Boguslawski, Paul A. Johnson, Peter A.Limacher, Matthew CHAN, Toon Verstraelen, Paul W. Ayers, *The Journal of Physical Chemistry A* **118**, 9058-9068 (2014).
- <sup>16</sup>John Bardeen, Leon N. Cooper, and J. Robert Schrieffer, *Physical Review* **108**, 1175 (1957).
- <sup>17</sup>Paul A. Johnson, Charles-Émile Fecteau, Frédéric Berthiaume, Samuel Cloutier, Laurie Carrier, Marianne Gratton, Patrick Bultinck, Stijn De Baerdemacker, Dimitri Van Neck, PeterLimacher, Paul W. Ayers, *The Journal of Chemical Physics* **153**, 104110 (2020).
- <sup>18</sup>P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck and P. Bultinck, *J. Chem. Theory and Computation* **9**, 1394-1401 (2013).
- <sup>19</sup>Peter A. Limacher, Taewon D. Kim, Paul W. Ayers, Paul A. Johnson, Stijn De

- Baerdemacker, Dimitri Van Neck, Patrick Bultinck, Mol. Phys. **112**, 853-862 (2014).
- <sup>20</sup>Katharina Boguslawski, Pawel Tecmer, Paul W. Ayers, Patrick Bultinck, Stijn De Baerdemacker, Dimitri Van Neck, Physical Review B **89**, 201106 (2014).
- <sup>21</sup>Katharina Boguslawski, Pawel Tecmer, Patrick Bultinck, Stijn De Baerdemacker, Dimitri Van Neck, Paul W. Ayers, Journal of Chemical Theory and Computation **10**, 4873-4882 (2014).
- <sup>22</sup>Katharina Boguslawski, Pawel Tecmer, Peter A. Limacher, Paul A. Johnson, Paul W. Ayers, Patrick Bultinck, Stijn De Baerdemacker, Dimitri Van Neck, The Journal of Chemical Physics **140**, 214114 (2014).
- <sup>23</sup>A.C. Hurley, J.E. Lennard-Jones and J.A. Pople, Proc. Roy. Soc. (London) **A220**, 446-455 (1953).
- <sup>24</sup>P. G. Lykos, R. G. Parr, J. Chem. Phys. **24**, 1166 (1956).
- <sup>25</sup>Peter G. Lykos, Robert G. Parr, J. Chem. Phys. **25**, 1301 (1956).
- <sup>26</sup>E. Kapuy, J. Chem. Phys. **44**, 956 (1966).
- <sup>27</sup>E. L. Mehler, K. Ruedenberg, D. M. Silver, J. Chem. Phys. **52**, 1181 (1970).
- <sup>28</sup>D. M. Silver, J. Chem. Phys. **52**, 299 (1970).
- <sup>29</sup>D. M. Silver, E. L. Mehler, K. Ruedenberg, J. Chem. Phys. **52**, 1174 (1970).
- <sup>30</sup>David M. Silver, Klaus Ruedenberg, Ernest L. MEHLER, J. Chem. Phys. **52**, 1206 (1970).
- <sup>31</sup>V. A. Rassolov, J. Chem. Phys. **117**, 5978 (2002).
- <sup>32</sup>Vitaly A. Rassolov, Feng Xu, Sophya Garashchuk, J. Chem. Phys. **120**, 10385 (2004).
- <sup>33</sup>Vitaly A. Rassolov, Feng Xu, J. Chem. Phys. **126**, 234112 (2007).
- <sup>34</sup>V. A. Rassolov, F. Xu J. Chem. Phys. **127**, 44104 (2007).
- <sup>35</sup>William A. Goddard III, Physical Review **157**, 81 (1967).
- <sup>36</sup>William A. Goddard III, Thom H. Dunning Jr., William J. Hunt, P. Jeffrey Hay Accounts of Chemical Research **6**, 11, 368-376 (1973).
- <sup>37</sup>Koushik Chatterjee, Katarzyna Pernal, Theor. Chim. Acta. **134**, 118 (2015).
- <sup>38</sup>A. John Coleman, *The AGP model for fermion systems*, International Journal of Quan-

- tum Chemistry **63**, 23-30 (1997).
- <sup>39</sup>E. Rosta and P.R. Surján, *Int. J. Quantum Chem.* **80**, 96-104 (2000).
- <sup>40</sup>E. Rosta and P.R. Surján, *J. Chem. Phys.* **116**, 878 (2002).
- <sup>41</sup>P.R. Surján, Ágnes Szabados, Péter Jeszenszki, Tamás Zoboki, *J. Math. Chem.* **50**, 534 (2012).
- <sup>42</sup>Péter Jeszenszki, Péter R. Nagy, Tamás Zoboki, Ágnes Szabados, Péter R. Surján, *Int. J. Quantum Chem.* **114**, 1048-1052 (2014).
- <sup>43</sup>B. A. Cagg, V. A. Rassolov, *J. Chem. Phys.* **141**, 164112 (2014).
- <sup>44</sup>Peter A. Limacher, Paul W. Ayers, Paul A. Johnson, Stijn De Baerdemacker, Dimitri Van Neck, Patrick Bultinck, *Physical Chemistry Chemical Physics* **16**, 5061-5065 (2014).
- <sup>45</sup>Dominic Földvári, Zsuzsanna Tóth, Péter R. Surján, Ágnes Szabados, *The Journal of Chemical Physics* **150**, 034103 (2019).
- <sup>46</sup>Katharina Boguslawski, Paul W. Ayers, *Journal of Chemical Theory and Computation* **11**, 5252-5261 (2015).
- <sup>47</sup>Tamás Zoboki, Ágnes Szabados, P.R. Surján, *Journal of Chemical Theory and Computation* **9**, 2602-2608 (2013).
- <sup>48</sup>Eric Neuscamman, *J. Chem. Phys.* **139**, 181101 (2013).
- <sup>49</sup>Eric Neuscamman, *J. Chem. Theory and Computation* **12**, 3149 (2016).
- <sup>50</sup>Eric Neuscamman, *Mol. Phys.* **114**, 577 (2016).
- <sup>51</sup>Eric Neuscamman, *Physical Review Letters* **109**, 203001 (2012).
- <sup>52</sup>Eric Neuscamman, *J. Chem. Phys.* **139**, 194105 (2013).
- <sup>53</sup>L. Zhao, E. Neuscamman, *J. Chem. Theory and Computation* **12**, 3719 (2016).
- <sup>54</sup>Airi Kawasaki, Osamu Sugino, *The Journal of Chemical Physics* **145**, 244110 (2016).
- <sup>55</sup>Thomas M. Henderson, Gustavo E. Scuseria, *The Journal of Chemical Physics* **151**, 051101 (2019).
- <sup>56</sup>T. M. Henderson and G. E. Scuseria, *J. Chem. Phys.* **153**, 084111 (2020).
- <sup>57</sup>A. Khamoshi, Guo P. Chen, Thomas M. Henderson, G. E Scuseria, *The Journal of*

- Chemical Physics **154**, 074113 (2021).
- <sup>58</sup>A. Khamoshi, F. A. Evangelista, G. E Scuseria, Correlating AGP on a quantum computer, Quantum Science and Technology **6**, 014004 (2020).
- <sup>59</sup>V. E. Elfving, M. Millaruelo, J. A. Gámez, C. Gogolin, Phys. Rev. **A103**, 032605 (2021).
- <sup>60</sup>A. Khamoshi, Thomas M. Henderson, G. E Scuseria, The Journal of Chemical Physics **151**, 184103 (2019).
- <sup>61</sup>Jean-David Moisset, Charles-Émile Fecteau, Paul A. Johnson, The Journal of Chemical Physics **156**, 214110 (2022).
- <sup>62</sup>E. K. Sklyanin, Letters in Mathematical Physics **47**, 275 (1999).
- <sup>63</sup>Patrick Cassam-Chenaï, *New ideas to reduce the computational complexity of non-orthogonal geminal methods for strongly-correlated electronic systems*, hal-03774583, <https://hal.archives-ouvertes.fr/hal-03774583>
- <sup>64</sup>Thomas Perez, Patrick Cassam-Chenaï, *Geometrical constraints to reduce complexity in quantum molecular systems*, hal-02531863, <https://hal.archives-ouvertes.fr/hal-02531863>
- <sup>65</sup>Thomas Perez, *Problèmes d'algèbre extérieure liés au calcul de fonctions d'ondes électroniques produits de géminales*, Université Côte d'Azur, 2020.  
<http://www.theses.fr/2020C0AZ4060>
- <sup>66</sup>D. Jayatilaka, D. J. Grimwood, A. Lee, A. Lemay, A. J. Russel, C. Taylor, S. K. Wolff, P. Cassam-Chenaï, A. Whitton, TONTO - A system for computational chemistry (2005).
- <sup>67</sup>A. Peruzzo et al., Nature Comm. **5**, 4213 (2014).
- <sup>68</sup>Thomas Perez, Patrick Cassam-Chenaï, Journal of Mathematical Chemistry **56**, 1428-1436 (2018).
- <sup>69</sup>Patrick Cassam-Chenaï, Frédéric Patras, Journal of Mathematical Physics **44**, 4884-4906 (2003).
- <sup>70</sup>G. Racah, Phys. Rev. **63**, 367 (1943).

- <sup>71</sup>D. J. Dean, M. Hjorth-Jensen, *Rev. Mod. Phys.* **75**, 607 (2003).
- <sup>72</sup>V. Zelevinsky and A. Volya, *Physics of atomic nuclei* **66**, 1781 (2003).
- <sup>73</sup>Laimutis Bytautas, Thomas M. Henderson, Carlos A. Jiménez-Hoyos, Jason K. Ellis, Gustavo E. Scuseria, *The Journal of Chemical Physics*, **135**, 044119 (2011).
- <sup>74</sup>R. G. Parr, F. O. Ellison, P. G. Lykos *J. Chem. Phys.* **24**, 1106 (1956).
- <sup>75</sup>R. McWeeny, *Proc. Roy. Soc. (London)* **A253**, 242 (1959).
- <sup>76</sup>Roy McWeeny, Yukiko Mizuno, *Proc. Roy. Soc. (London)* **A259**, 554-577 (1961).
- <sup>77</sup>S. Wilson, *J. Chem. Phys.* **64**, 1692 (1976).

**Ground state energies from 2D-block,  $\{G_\theta, \sigma_x\}$ -type calculations  
at geometries derived from experiment (in Hartree)**

Homonuclear diatomics	N <sub>2</sub>	N <sub>2</sub>	O <sub>2</sub>	F <sub>2</sub>
$r_e^{exp}$ ( $\text{\AA}$ ) <sup>a</sup>	1.098	1.098	1.2075	1.412
Basis set <sup>b</sup>	cc-pVTZ	cc-pVDZ	cc-pVDZ	cc-pVDZ
$E_{\{G_\theta, \sigma_x\}}^0$ <sup>c</sup>	-109.134390	-109.082362	-149.718718	-198.853436
$E_{APSG}^0$ <sup>d</sup>	-109.110630	-109.058601	-149.689436	-198.840387
$E_{RHF}^0$	-108.983412	-108.954087	-149.542930	-198.685664
nb. of geminals	7	7	8	9
nb. of blocks	30	14	14	14
nb. of parameters	70/90	35/42	36/42	42/42
nb. terms in overlap <sup>e</sup>	64974	360	343	1116
cpu time (s) <sup>f</sup>	0.4	0.01	0.01	0.02

Table I. Ground state energies (in Hartree) from 2D-block,  $\{G_\theta, \sigma_x\}$ -type calculations at internuclear distances,  $r_e^{exp}$  (in  $\text{\AA}$ ), derived from experiment, for second row homonuclear diatomics. The 2D-block,  $\{G_\theta, \sigma_x\}$ -type energies  $E_{\{G_\theta, \sigma_x\}}^0$ , can only be considered as upper bounds since the orbital grouping into 2D-blocks and the  $\sigma_x$ -type allocation to the different geminals have not been systematically investigated (see main text for details). Inputs and output files corresponding to these calculations are available in the "tests/geminal" sub-directory of the "tonto/Nice-branch" repository. The line "nb. of parameters" gives the number of parameters actually used over the maximum possible within the model.

<sup>a</sup> From NIST website: <https://cccbdb.nist.gov/>.

<sup>b</sup> From: T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

<sup>c</sup> Orbital grouping and  $\sigma_x$ -type allocation have not been systematically investigated.

<sup>d</sup> From the SSG code of V. A. Rassolov in QCHEM: Y. Shao et al., Mol. Phys. **113**, 184 (2015).

<sup>e</sup> For the block types allocation that gives the  $E_{\{G_\theta, \sigma_x\}}^0$  energy reported.

<sup>f</sup> Sequential calculations, on a laptop with Intel® Core™ i9-10885H CPU @ 2.40GHz.