# The Electronic Mean-Field Configuration Interaction method: III- the p-orthogonality constraint

Patrick Cassam-Chenaï, Vitaly Rassolov

# The Electronic Mean-Field Configuration Interaction method: III- the $p$-orthogonality constraint

Patrick Cassam-Chenaï

*Laboratoire J. A. Dieudonné, UMR 6621 du CNRS, Faculté des Sciences, Parc Valrose, 06108 Nice cedex 2, France. cassam@unice.fr*

Vitaly Rassolov

*Department of Chemistry and Biochemistry, University of South Carolina, 631 Sumter Street, Columbia, South Carolina 29208.*

**Abstract**

The concept of $p$-orthogonality between electronic states, which generalizes common orthogonality and strong orthogonality, provides a natural hierarchy for group function methods such as the electronic mean field configuration interaction method. In this letter, this theoretical concept is applied in numerical calculations for the first time. The accuracy of the geminal mean field configuration interaction wave functions of simple molecular systems is studied as the orthogonality constraint between the geminals relaxes from 1-orthogonality (that is strong orthogonality) to 2-orthogonality, to no orthogonality constraint at all.

*Preprint submitted to Chemical Physics Letters*

## 1 Introduction

The concept of $p$-orthogonality between electronic wave functions has been proposed for the first time by S. Wilson [1]. One of the authors has rediscovered independently this concept and has defined it in the more general context of quantum states, either pure or mixed (i.e. ensemble) states, of systems of identical particles, either fermions or bosons [2].

This letter presents the application of this concept to the simplest, non trivial, electronic mean field configuration interaction (EMFCI) method that is the geminal mean field configuration interaction (GMFCI) method or when iterated, the geminal self consistent field (GSCF) method [3]. Our goal is to assess on simple atomic and molecular systems for which full configuration interaction (CI) references are available, how the accuracy of the GMFCI or GSCF methods is affected by imposing the $p$-orthogonality constraint for $p \in \{1, 2\}$. To the best of our knowledge, it is the first numerical application of $p$-orthogonality as such.

At the same time, the present study complements our previous work [4] (part II of this series of papers) on guess geminals and demonstrates that the strongly orthogonal geminals with variationally optimized Hilbert subspaces [5] (called "antisymmetrized product of singlet-type strongly orthogonal geminals", APSSG, by Rassolov et al. or "restricted singlet-type strongly orthogonal geminals", RSSG, in the case of spin restricted calculations) is an extremely good guess for the EMFCI method.

Throughout the letter, we use the following notation: lower case greek symbols $\psi_i$ are one-electron spinorbitals, upper case $\Psi_i$ are antisymmetric functions of one or more electrons and are called "electron group functions" or just "group functions". The antisymmetrized product of electron group functions $\Psi = \mathcal{A}[\Psi_1(r_1, \cdots, r_{n_1}), \cdots, \Psi_k(r_{n-n_k+1}, \cdots, r_n)]$, where $\mathcal{A}$ is the antisymmetrizing operator, (that is the normalized sum of signed permutations permuting the variables $r_j$ among the different group functions), $n_i$ the number of electrons in $\Psi_i$, and $n = \sum_i n_i$ the total number of electrons, is denoted with the

Grassmann product ($\wedge$) as

$$\Psi = \Psi_1 \wedge \cdots \wedge \Psi_k. \tag{1}$$

An introduction to the Grassmann product formalism can be found in Ref. [9,10]. Note that in the literature, the term "group function" is usually employed for an antisymmetrical product function $\Psi$ rather than for the $\Psi_i$'s.

The letter is organised in the following way: We first recall the definition of $p$-orthogonality, then we briefly introduce the EMFCI and GSCF methods with or without $p$-orthogonality imposed between group functions. Next, we illustrate numerically the effect of this constraint on a few chosen examples. Finally, we conclude on the prospect of tuning the $p$-orthogonality constraint to describe accurately the physics of a molecular system at the least computational cost within a group function frame.

## 2 $p$-orthogonality

An easy way to introduce $p$-orthogonality between two electronic states consists in considering the eigenfunctions associated with the non zero eigenvalues of their respective $p^{th}$-order reduced density operators. Let us call these functions the "populated, natural, $p$-electron functions", each set spans a vector space that we have termed the "$p$-internal space" of the related electronic state. Another way to look at the $p$-internal space of an electronic pure state which justifies our terminology is the following characterization [10]: a function, $\Phi$, belongs to the $p$-internal space of an $n$-electron wave function if and only if there exists an $(n-p)$-electron function such that its annihilation in the $n$-electron wave function (in the second quantization sense) gives $\Phi$. In other words, the $p$-internal space of an $n$-electron wave function is the vector space of all the $p$-electron functions obtained by annihilation of an $(n-p)$-electron function in the $n$-electron wave function.

For example, the 1-internal space, or simply "internal space" of a wave function $\Psi$, denoted by, $\mathcal{I}^1[\Psi]$, is the space spanned by the occupied, natural spinorbitals. The 2-

internal space, denoted by, $\mathcal{I}^2[\Psi]$, is the space spanned by the occupied, natural geminals. The $n$-internal space is the one-dimensional vector space spanned by the wave function $\Psi$, $\mathbb{C} \cdot \Psi$. The $p$-internal space of a single configuration function (Slater determinant) built over a set of $n$ orthogonal spinorbitals, $\Psi := \psi_1 \wedge \ldots \wedge \psi_n$, is the $\binom{n}{p}$-dimensional vector space spanned by the $p$-particle functions, $\psi_{i_1} \wedge \ldots \wedge \psi_{i_p}$, built over $p$ spinorbitals of $\Psi$.

Then, two electronic states of, say, $n_1$ and $n_2$ electrons respectively, are said $p$-orthogonal with $1 \leq p \leq inf(n_1, n_2)$, if and only if their $p$-internal spaces are orthogonal. Let us consider two pure states represented by wave functions $\Psi_1$ and $\Psi_2$, respectively. We see immediately that if $n_1 = n_2 = n$, $n$-orthogonality is the usual orthogonality between wave functions, since $\mathcal{I}^n[\Psi_i] = \mathbb{C} \cdot \Psi_i$, $\forall i \in \{1, 2\}$, so that

$$\mathcal{I}^n[\Psi_1] \perp \mathcal{I}^n[\Psi_2] \Longleftrightarrow \Psi_1 \perp \Psi_2 \Longleftrightarrow \langle \Psi_1 | \Psi_2 \rangle = 0.$$

At the other end, 1-orthogonality between $\Psi_1$ and $\Psi_2$ amounts to strong orthogonality [6–8].

An important property is that $p$-orthogonality implies $q$-orthogonality for all $q \geq p$. So, $p$-orthogonality provides us with a graded orthogonality concept for electronic states. It discriminates between pairs of wave functions that are all orthogonal in the traditional sense: Let $\psi_1, \ldots, \psi_{2n}$ be $2n$ orthonormal spinorbitals. For $p > 0$, the Slater determinants $\Psi_1 := \psi_1 \wedge \ldots \wedge \psi_n$ and $\Psi_2 := \psi_1 \wedge \ldots \wedge \psi_{n-p} \wedge \psi_{n+1} \ldots \wedge \psi_{n+p}$ are $(n-p+1)$-orthogonal but not $(n-p)$-orthogonal. Furthermore, the concept is well-defined for multiconfigurational functions of different numbers of electrons: The pairs $\Psi_1 := \psi_1 \wedge \psi_2 \wedge \psi_3 + \psi_4 \wedge \psi_5 \wedge \psi_6$ and $\Psi_2 := \psi_1 \wedge \psi_7 + \psi_2 \wedge \psi_8$ are 2-orthogonal (it is impossible to obtain a function not orthogonal to $\Psi_2$ by annihilating a spinorbital in $\Psi_1$) but not 1-orthogonal since both $\psi_1$ and $\psi_2$ belongs to their respective internal spaces.

A remarkable fact that justifies the relevance of $p$-orthogonality to quantum chemistry, is that this purely geometrical concept is related in a neat and straightforward manner to the expression of matrix elements of quantum observables between antisymmetrized

products of electron group functions [1,2]. The stronger the $p$-orthogonality between the factors of the antisymmetrized products, the simpler the combinatorics involved in the matrix element expressions, and the lower their computational cost. So, imposing $p$-orthogonality constraints to antisymmetrized products of electron group function for increasing values of $p$, naturally leads one to consider a corresponding hierarchy of group function approximation methods.

## 3 $p$-orthogonally constrained EMFCI methods

In quantum chemistry, variational methods to solve the electronic Schrödinger equation are usually classified into two classes: those which use orthogonal orbitals and those which use non-orthogonal orbitals. In the first class, the most general is arguably the ORMAS (occupation-restricted-multiple active spaces) method [11]. ORMAS-CI and ORMAS-SCF encompass the traditional CI methods and multi-configuration self-consistent field (MC-SCF) methods. In the second class, one finds methods based on valence-bond wave functions, such as those implemented in the code TURTLE [12], and XMVB [13], or on more general group function products as implemented in the code VB2000 [14]. A fairly extensive bibiliography on early works using group function product wave functions has been given in a previous article of this series [3], a few other relevant references are [15–22]. More references on non-orthogonal methods can be found in Ref. [23]. Note also a recently proposed method [24], which implements the idea of finding an optimal wave function of a fixed length (see Remark 4.10 of Ref. [9]) that is the best sum of a fixed number of Slater determinants with no orthogonality nor normalization restriction on their spinorbitals.

An Electronic Mean Field Configuration Interaction (EMFCI) calculation is an $n$-electron configuration interaction (CI) calculation, where the set of orthogonal basis functions considered has a group function product structure of the form: $(\Psi_1^0 \wedge \cdots \wedge \Psi_j^k \wedge \cdots \wedge \Psi_r^0)_{0\leq k\leq N}$. It can be regarded as a configuration interaction for the "active" group $j$ in the mean field of the $(r-1)$ "spectator" electron group wave functions $\Psi_i^0$ for $i \neq j$.

An Electronic Self-Consistent Field Configuration Interaction (ESCFCI) calculation for a fixed set of electron numbers, $(n_i)_{1\leq i\leq r}$, is defined to be the variational method that finds the "best" (with respect to the variational principle) $n$-electron wave functions of the form given by Eq. (1). The best ground state can be obtained in principle by successive EMFCI steps, provided that, at a given step where group $j$ is active, each spectator state $\Psi_i^0$ (for $i \neq j$) is chosen to be the group function corresponding to the ground state of the last EMFCI step where group $i$ was active.

Instead of trying to achieve self-consistency with respect to the variation of the original group functions, one can choose to perform an EMFCI step for another set of integers, $(n'_i)_{1\leq i\leq r'}$ with $r' < r$, corresponding to a coarser partition of the electronic system, that is a partition into larger electron groups such that each old group is wholly included into one of the new groups.

In the original form of the EMFCI and ESCFCI methods, no orthogonality constraint is imposed on the $\Psi_i$. In particular, in the simple case where for all $i$, $n_i = 2$, both antisymmetrized product of strongly orthogonal geminals (APSG) (i.e. 1-orthogonal group functions) and antisymmetrized geminal product (AGP) of extreme type, $\Psi_1 \wedge \cdots \wedge \Psi_1$, (where group functions are all equal, so not even 2-orthogonal), are considered by the variational process. In this work, we study this simple case, where all the group functions are geminals, with different $p$-orthogonality constraints.

We will consider two cases of $p$-orthogonality constraints:
- case 1: the active group, say group 1 without loss of generality, is $p$-orthogonal to every spectator group functions, $\Psi_i^0,\ i > 1$.
- case 2: the active group is $p$-orthogonal to every spectator group functions, and every product of spectator group functions, $\Psi_{i_1}^0 \wedge \cdots \wedge \Psi_{i_k}^0,\ \forall k \in \{1, \cdots, r-1\}$. Note that, for geminal product wave functions and 2-orthogonality, it is sufficient to impose 2-orthogonality of the active group to every spectator group functions, $\Psi_i^0,\ i > 1$, and every product of pairs of spectator group functions, $\Psi_{i_1}^0 \wedge \Psi_{i_2}^0,\ i_1, i_2 > 1$, see Appendix A.

In the case $p = 1$, if the spectator group functions $\Psi_i^0$ are themselves 1-orthogonal to one another, the two cases are equivalent. However, in general, constraint 2 that we call spectator product $p$-orthogonality (SP$p$-O) is a stronger requirement than constraint 1 called spectator function $p$-orthogonality (SF$p$-O).

Imposing the $p$-orthogonality constraint is close in spirit with the partial release of the antisymmetry constraint in spin-coupled wave functions [25–27]. It is well known that there is no need to antisymmetrize two group functions that are 1-orthogonal. More generally, energy expressions with antisymmetry limited to $(p-1)$ interchanges between spin-coupled group functions (that is to say, expansions obtained by neglecting interchanges of more than $(p-1)$ pairs of particles) are related to $p$-orthogonality in the sense that they are exact for fully antisymmetrized functions built from $p$-orthogonal group functions. However, our goal here, is not to optimize a general wave function with an approximate energy expression, but to optimize a $p$-orthogonal constrained wave function with an exact energy expression.

In the following, the effect of SP$p$-O and SF$p$-O, with $p \in \{1, 2\}$, are evaluated for the GMFCI and GSCF methods on some singlet molecules. We have limited our study to molecular systems whose size is small enough so as to allow one to perform a full CI calculation that can serve as a reference.

### *3.1 p-orthogonally constrained GMFCI*

We have used linear chains of equidistant H nuclei to test the SP1-O $\equiv$ SF1-O, SP2-O and SF2-O constraints on the GMFCI method. These systems are of interest in connection to metallic hydrogen and are often used for benchmarking purposes [28]. The internuclei distance between adjacent nuclei was set to 1 angström for all the systems, which is close to the value one usually obtains by optimizing their geometry within the constraint of equidistant H nuclei. GMFCI calculations were performed for different chain lengths and basis sets, see Tab. 1. Results were found fairly insensitive to the value of linear

dependency threshold (see Appendix B) from $ld = 10^{-2}$ to $ld = 10^{-5}$, however, they are less accurate for $ld = 10^{-1}$ and numerical instabilities may occur for $ld < 10^{-5}$ (or $ld < 10^{-4}$ for the largest systems).

Two types of one-orthogonal geminal product functions were considered. The first one, (row "step 0/ CHFO") corresponds to the best GMFCI solution starting from guess, singlet geminals built as paired spin-$\alpha$, spin-$\beta$ canonical Hartree-Fock orbitals. By "best" we mean that the choice of the active group is the one that gives the lowest ground state energy. More explicitly, an RHF calculation is first performed, and a wave function of the form, $\Psi_{HF} = \psi_1^\alpha \wedge \psi_1^\beta \wedge \cdots \wedge \psi_r^\alpha \wedge \psi_r^\beta$, (where the superscript $\alpha$ or $\beta$ indicates the spin of the HF spinorbitals), is obtained. Then, the $r$ possible GMFCI calculations with $r-1$ canonical HF spectator geminals are performed and the wave function of the form $\Psi_{step0/CHFO} = \psi_1^\alpha \wedge \psi_1^\beta \wedge \cdots \wedge g_i \wedge \cdots \wedge \psi_r^\alpha \wedge \psi_r^\beta$ giving the lowest ground state energy is selected. The corresponding energy is tabulated in row "step 0/CHFO" of Tab. 1. Here $g_i := \sum_{b_1,b_2=1}^{m} c_{b_1,b_2}^i \; \psi_{b_1}^\alpha \wedge \psi_{b_2}^\beta$ is the geminal obtained by a CI calculation in an orthonormal basis set of the linear space spanned by the functions of the forms $\psi_1^\alpha \wedge \psi_1^\beta \wedge \cdots \wedge \psi_j^\alpha \wedge \psi_j^\beta \wedge \cdots \wedge \psi_r^\alpha \wedge \psi_r^\beta$ and $\psi_1^\alpha \wedge \psi_1^\beta \wedge \cdots \wedge \left(\psi_j^\alpha \wedge \psi_k^\beta + \psi_k^\alpha \wedge \psi_j^\beta\right) \wedge \cdots \wedge \psi_r^\alpha \wedge \psi_r^\beta$. Note that strong orthogonality is automatically enforced in such a GMFCI because any geminal for the active group having an internal spectator spinorbital would give an $r$ geminal product configuration that is zero, since the spectator product function is a Slater determinant.

The other type of one-orthogonal geminal product function referred to as SP1-O in Tab. 1, is the best possible singlet geminal product function for a given orbital basis set. By "best" we mean that both the partitioning of the Hilbert space into $r$ orthogonal Hilbert subspaces, (sometimes called Arai spaces [29]), and the geminal coefficients have been optimized to minimize the ground state energy. In other words, it is the RSSG solution thoroughly studied by one of us [5,30].

Both types of one-orthogonal geminal product functions can serve as starting guess for further SP2-O, SF2-O or unconstrained GMFCI. The comparison of rows "$E^0_{step0/CHFOs}$"

and "$E^0_{SP1-O}$" of Tab. 1 demonstrates clearly the superiority of the RSSGs energies. However, when no constraint is enforced and GMFCI calculations are iterated until self-consistency, the same energy should be obtained regardless of the starting guess. The question is whether the difference $E^0_{step0/CHFOs} - E^0_{SP1-O}$ is caught up in a single GMFCI step or only after many such steps. The situation is even less clear with orthogonality constraints, in particular, the dimension of the 2-internal space of an RSSG spectator product wavefunction is in general greater than its best step 0/CHFO counterpart, so we could expect SP2-orthogonality to be more stringent for RSSG than for best step 0/CHFO guess at the first GMFCI step. In fact, SP2-O, SF2-O or unconstrained GMFCI calculations starting with $\Psi_{step0/CHFO}$ not reported here, show that the magnitudes of the improvements with respect to $E^0_{step0/CHFOs}$ are of the same order as those with respect to $E^0_{SP1-O}$ reported in rows "$E^0_{SP2-O}$", "$E^0_{SF2-O}$" and "$E^0_{GMFCI}$" of Tab. 1, which correspond to SP2-O, SF2-O or unconstrained GMFCI calculations starting with $\Psi_{SP1-O}$. So, the advantage of the RSSG wave function is preserved after a GMFCI calculation.

The discrepancy of GMFCI with respect to Full CI in the STO-3G basis set grows linearly with the number of electron pairs, $r$, according to, $\Delta E \simeq 16 * r - 10mH$, for $r = 2, 3, 4$. In contrast, it seems insensitive to the size of the basis set for $H_6$. In all calculations, the SF-2-orthogonality constraint was found essentially transparent. For STO-3G calculations, the configurations it would eliminate are already eliminated by the removal of quasi-linear dependencies. For larger basis sets, it does serve to eliminate two CI coefficients from the GMFCI expansion, while the final energy is not affected. Note for comparison that the GSCF energy converges towards $-3.231555$ Hartree for the H6 STO-3G calculation, (with no guarantee that it is the absolute minimum).

The SP-2-orthogonality constraint not only simplifies the matrix element calculation but also reduces the number of configurations, as can be seen from the "SP2-O gem. nb." row of Tab. 1. However, the ground state energies obtained are the same as the unconstrained calculations to sub-mHartree precision.

### *3.2 p-orthogonally constrained GSCF*

The effect of orthogonality constraints on the GSCF method is illustrated in Tab. 2. The reported GSCF energies are upper bounds of the exact ones after a large number of iterations: since second derivatives are not calculated, it is not possible to ascertain that the algorithm reaches a minimum, not to mention the absolute minimum. Note also that SP2-orthogonality is not investigated because it is not preserved in the GSCF algorithm implemented in this work which just consists in iterating GMFCI calculations: Even if we start with a function $\Psi_1^0 \wedge \cdots \wedge \Psi_j^0 \wedge \cdots \wedge \Psi_r^0$ such that each $\Psi_j^0$ is 2-orthogonal to every other geminal $\Psi_i^0$, and to every geminal pair product, $\Psi_i^0 \wedge \Psi_k^0$, $\forall i, k \neq j$, (see Appendix A); after a SP2-orthogonality constrained GMFCI step, where without loss of generality we can assume that the active group is group 1, the resulting wave function, $\Psi_1^{GMFCI} \wedge \Psi_2^0 \wedge \cdots \wedge \Psi_r^0$ does not necessarily satisfy the same property. By construction, $\Psi_1^{GMFCI}$ will be 2-orthogonal to every spectator geminal product, but for instance, $\Psi_2^0$ has no reason be 2-orthogonal to, say, $\Psi_1^{GMFCI} \wedge \Psi_3^0$. So, SP2-orthogonality can still be enforced in successive GMFCI steps and usefully reduce the number of configurations, but unfortunately, it cannot be taken advantage of to reduce significantly the matrix element computation effort.

Tab. 2 shows again that RSSG ground state energies are significantly lower than those obtained from best step 0/CHFO GMFCI calculations for the last three molecules. Although we could expect that after many iterations both guesses would lead to the same GSCF energy, in practice it only happens for the simplest systems. For larger systems, we have often observed that when starting from best step 0/CHFO GMFCI calculations, convergence is stalled before reaching the value obtained from calculations using a RSSG starting guess. This confirm the superiority of the latter. The SF2-orthogonal calculations remain remarkably close to the unconstrained result for GSCF, although not identical as for a single GMFCI step. The GSCF are themselves remarkably close to the full CI results which tends to prove that the absolute minima have been found although this cannot be ascertained. The difference between GSCF and Full CI, which

is due to the neglect in GSCF of the correlated excitations between geminal functions, gives an idea of the weight of the latter in the wave funcion.

Besides the total energy, we have investigated the electric dipole moment for two systems of Tab.2, namely LiH and BH, see Tab. 3. The excellent agreement with the full CI values confirms the validity of our geminal product functions.

## 4 Conclusion

Standard guess geminals constructed from the HF canonical orbitals and used in previous publications have been found clearly inferior to RSSG geminals, obtained by optimising strongly orthogonal Hilbert subspaces for the different electron groups, in ground state energy calculations. However, the ideas developed in [4] to rotate HF canonical orbitals in order to improve excited electronic state calculations can be transposed to RSSG optimized orbitals. Within each optimized Hilbert subspaces, orbitals can be rotated to minimise the low lying excited states obtained by single excitation from the subspace. The combination of such a technique with RSSG geminals would provide optimal guess geminals for both ground and excited states.

The RSSG calculation providing guess geminals can be regarded as a GSCF calculation with SF1-orthogonality enforced. Unconstrained GSCF calculations are very accurate, and the SF2-orthogonality constraint is virtually transparent. However the present algorithm that performs GSCF as an iterated GMFCI with no truncation applied to the geminal basis set converges poorly, arguably because the geminals in a Grassmann product function are strongly inter-related and should be optimized simultaneously rather than one at a time as in the present algorithm. The SF2-orthogonality constraint hardly remedies the situation. The SP2-orthogonality constraint is much more efficient in this respect, but has not been studied here because it is lost after the first iteration with the present algorithm.

In fact, if one wants to adhere to the electronic mean field CI philosophy, one should aim at performing only a few GMFCI steps and play with basis set truncation thresholds both to prevent excited states from loosing their physical relevance and to save CPU time. Then, even a single GMFCI step improves greatly over the HF calculation. Moreover, starting from an SF1-orthogonality constrained GSCF (that is RSSG), and performing a single GMFCI step with or without an SF2-orthogonality constraint, one can obtain an accurate wave function with released strong orthogonality restriction. Such a calculation will be affordable even for large systems, if one uses basis set truncation thresholds to limit the configuration space, (as used in part II of this series), and/or if one uses the SP2-orthogonality constraint, which simplifies the Hamiltonian and overlap matrix calculation, and which has been shown to preserve the GMFCI accuracy to a satisfactory level.

However, we are currently investigating a form of the SP2-orthogonality constraint, that is invariant under permutation of group functions and passes through successive GMFCI steps, which might remedy the GSCF convergence problem. In addition, the same permutation-invariant SP2-orthogonality constraint passes through recursion steps in matrix element calculations. This gives us hope that polynomial scaling with system size in the computational cost of a post strongly-orthogonal ansatz can be achieved.

More generally, we hope that future works will demonstrate that EMFCI calculations with $p$-orthogonality constraint schemes, adapted to a hierarchy of electron groups in a molecule, in a manner reminiscent of [31] but with a partioning based on optimized Arai spaces rather than orbitals, provide a good compromise between accuracy and computational cost.

## Acknowledgements

We acknowledge the french ANR for fundings (Projet AHBE). We are grateful to Prof. P. R. Surjàn for drawing our attention to the work of S. Wilson and to Dr. F. Patras

and Prof. P. Diaconis for fruitful discussions. GMFCI, and GSCF computations were performed locally and at the SIGAMM Mesocentre with the code TONTO-CHEM distributed on SourceForge. We acknowledge the referees for suggesting many improvements to our manuscript.

## References


[1] S. Wilson, J. Chem. Phys 64 (1976) 1692.

[2] P. Cassam-Chenaï, Phys. Rev. A77 (2008) 032103.

[3] P. Cassam-Chenaï, J. Chem. Phys 124 (2006) 194109-194123.

[4] P. Cassam-Chenaï, G. Granucci Chem. Phys. Lett. 450 (2007) 151-155.

[5] V. A. Rassolov, F. Xu, J. Chem. Phys 127 (2007) 44104. and references therein

[6] A.C. Hurley, J.E. Lennard-Jones and J.A. Pople, Proc. Roy. Soc. (London) A220 (1953) 446-455.

[7] R. G. Parr, F. O. Ellison, P. G. Lykos J. Chem. Phys 24 (1956) 1106.

[8] R. McWeeny, Proc. Roy. Soc. (London) A253 (1959) 242.

[9] P. Cassam-Chenaï, J. Math. Chem. 15 (1994) 303.

[10] P. Cassam-Chenaï, F. Patras, J. Math. Phys. 44 (2003) 4884-4906.

[11] J. Ivanic, J. Chem. Phys 119 (2003) 9364.

[12] J. Verbeek, J.H. Langenberg, C.P. Byrman, F. Dijstra, J.H. van Lenthe, TURTLE, an ab initio VB/VBSCF program. Utrecht, 1993.

[13] L. Song, Y. Mo, Q. Zhang, W. Wu, J. Comp. Chem. 26 (2005) 51492.

[14] J. Li, R. McWeeny, Int. J. Quantum Chem. 89 (2002) 208.

[15] S. Bratož, P. Durand, J. Chem. Phys 43 (1965) 2670.

[16] L. Szasz, J. Chem. Phys 49 (1968) 679.

[17] K. J. Miller, K. Ruedenberg, J. Chem. Phys 48 (1968) 3444, ibid p.3450.

[18] K. J. Miller, K. Ruedenberg, J. Chem. Phys 48 (1968) 3450.

[19] I. Ts. Lyast, T. A. Yampolskaya, É. M. Zhdanova, Zhurnal Strukturnoi Khimii 10 (1969) 757-758.

[20] A. M.Tokmachev, R. Dronskowski, J. Comp. Chem. 27 (2005) 296-308.

[21] A. M.Tokmachev, A. L. Tchougreff, Int. J. Quantum Chem. 106 (2006) 571-587.

[22] A. M.Tokmachev, R. Dronskowski, Chem. Phys. 322 (2006) 423-432.

[23] F. E Penotti, J. Comp. Chem. 27 (2006) 749.

[24] G. Beylkin, M. J. Mohlenkamp, F. Prez, J. Math. Phys. 49 (2008) 032107.

[25] P. D. Dacre, R. McWeeny, Proc. Roy. Soc. A 317 (1970) 435-454.

[26] A. Howard, `Ph. D. thesis`, University of York, (1979).

[27] D. J. Klein, C. H. Carlisle, F. A. Matsen, *Adv. Quantum Chem.* 5 (1970) 219.

[28] D. A. Mazziotti, Phys. Rev. Let. 93 (2004) 213001.

[29] T. Arai, J. Chem. Phys 33 (1960) 95.

[30] V. A. Rassolov, J. Chem. Phys 117 (2002) 5978.

[31] S. Wilson, Int. J. Quantum Chem. 74 (1999) 135.

## Appendix A: SP2-O for geminal product functions

Here we prove that, in the case of a GMFCI calculation, it is sufficient to impose 2-orthogonality of the active group to every spectator group functions and every product of pairs of spectator group functions, to enforce the SP2-O constraint, that is to say, to obtain an active group that is 2-orthogonal to every product of spectator group functions, so, in particular, to the product of all the $(r-1)$ spectator group functions in the case of an $r$ electron pair system.

Lemma:
Assume that $\forall i,j>1$, $\langle\Psi_i|\Psi_1\rangle=0$ and $\Psi_1$ is 2-orthogonal to $\Psi_i\wedge\Psi_j$, where $\Psi_1,\ldots,\Psi_r$ are $r$ geminal functions then $\Psi_1$ is 2-orthogonal to $\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}$, $\forall\{k_1,\cdots,k_s\}\subseteq\{2,\cdots,r\}$.

Proof: One has to show that $\Psi_1$ is an external geminal of $\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}$, that is to say, in the second quantization language that its annihilation in $|\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}\rangle$ is zero, or equivalently, in the exterior algebra formalism [9], that its interior product, denoted by the symbol $\hookleftarrow$, is zero: $\Psi_1\hookleftarrow\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}=0$. For any $2(s-1)$-electron wave function, $\Phi$, let us compute $\langle\Psi_1\hookleftarrow\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}|\Phi\rangle$. The Hopf algebra formalism allows us to perform mechanically such a calculation (see part I of this series [3] for notation):

$$
\begin{aligned}
&\langle\Psi_1\hookleftarrow\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}|\Phi\rangle=\langle\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}|\Psi_1\wedge\Phi\rangle\\
&=\langle\mathcal{X}^{[s-1]}(\Psi_{k_1}\otimes\cdots\otimes\Psi_{k_s})|\mathcal{X}(\Psi_1\otimes\Phi)\rangle\\
&=\langle\mathcal{Y}\circ\mathcal{X}^{[s-1]}(\Psi_{k_1}\otimes\cdots\otimes\Psi_{k_s})|\Psi_1\otimes\Phi\rangle\\
&=\langle(\mathcal{X}^{[s-1]}\otimes\mathcal{X}^{[s-1]})\circ T^{(2,s)}\circ(\mathcal{Y}(\Psi_{k_1})\otimes\cdots\otimes\mathcal{Y}(\Psi_{k_s}))|\Psi_1\otimes\Phi\rangle\\
&=\langle(\mathcal{X}^{[s-1]}\otimes\mathcal{X}^{[s-1]})\circ T^{(2,s)}\circ\left(\sum_{i\in\{k_1,\cdots,k_s\}}\mathcal{Y}_{0,2}(\Psi_{k_1})\otimes\cdots\otimes\mathcal{Y}_{2,0}(\Psi_i)\otimes\cdots\otimes\mathcal{Y}_{0,2}(\Psi_{k_s})\right.\\
&\left.+\sum_{i<j\in\{k_1,\cdots,k_s\}}\mathcal{Y}_{0,2}(\Psi_{k_1})\otimes\cdots\otimes\mathcal{Y}_{1,1}(\Psi_i)\otimes\cdots\otimes\mathcal{Y}_{1,1}(\Psi_j)\otimes\cdots\otimes\mathcal{Y}_{0,2}(\Psi_{k_s})\right)|\Psi_1\otimes\Phi\rangle.
\end{aligned}
$$

Introducing the "hat" notation: $(\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s})_{\hat{i}}$ means that $\Psi_i$ is taken out of the product $\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s}$, and $(\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s})_{\hat{i},\hat{j}}:=((\Psi_{k_1}\wedge\cdots\wedge\Psi_{k_s})_{\hat{i}})_{\hat{j}}$, we obtain,

$$\langle \Psi_1 \hookleftarrow \Psi_{k_1} \wedge \cdots \wedge \Psi_{k_s} | \Phi \rangle = \sum_{i \in \{k_1, \cdots, k_s\}} \langle \Psi_i | \Psi_1 \rangle \langle (\Psi_{k_1} \wedge \cdots \wedge \Psi_{k_s})_{\hat{i}} | \Phi \rangle \quad -$$
$$\sum_{i<j \in \{k_1, \cdots, k_s\}} \sum_{I,J \in \mathcal{P}_{2,1}} \langle \mathcal{Y}_{1,1}(\Psi_i)_I \wedge \mathcal{Y}_{1,1}(\Psi_j)_J | \Psi_1 \rangle \langle (\Psi_{k_1} \wedge \cdots \wedge \Psi_{k_s})_{\hat{i},\hat{j}} \wedge \mathcal{Y}_{1,1}(\Psi_i)_{\bar{I}} \wedge \mathcal{Y}_{1,1}(\Psi_j)_{\bar{J}} | \Phi \rangle. \tag{2}$$

The first hypothesis of the lemma ensures that the first part of the right-hand-side of Eq.(2) is zero. So, this equation simplifies, and it gives for $s = 2$ and any geminal $\Phi$,

$$\langle \Psi_1 \hookleftarrow \Psi_i \wedge \Psi_j | \Phi \rangle = \sum_{I,J \in \mathcal{P}_{2,1}} \langle \mathcal{Y}_{1,1}(\Psi_i)_I \wedge \mathcal{Y}_{1,1}(\Psi_j)_J | \Psi_1 \rangle \langle \mathcal{Y}_{1,1}(\Psi_i)_{\bar{I}} \wedge \mathcal{Y}_{1,1}(\Psi_j)_{\bar{J}} | \Phi \rangle. \tag{3}$$

However, this expression is zero according to the second hypothesis of the lemma for any $\Phi$, so we obtain,

$$\sum_{I,J \in \mathcal{P}_{2,1}} \langle \mathcal{Y}_{1,1}(\Psi_i)_I \wedge \mathcal{Y}_{1,1}(\Psi_j)_J | \Psi_1 \rangle \mathcal{Y}_{1,1}(\Psi_i)_{\bar{I}} \wedge \mathcal{Y}_{1,1}(\Psi_j)_{\bar{J}} = 0. \tag{4}$$

For any $i < j \in \{k_1, \cdots, k_s\}$, Eq.(4) can be bracketed by $\langle (\Psi_{k_1} \wedge \cdots \wedge \Psi_{k_s})_{\hat{i},\hat{j}} \wedge$ on the left and by an arbitrary $2(s-1)$-electron wave function, $|\Phi\rangle$, on the right, it will still be zero. It suffices to sum up over all $i < j \in \{k_1, \cdots, k_s\}$ to recover the second part of the right-hand-side of Eq.(2), therefore, the latter is zero. Since we have seen that the first part of the right-hand-side of Eq.(2) is also zero for all $\Phi$, we conclude that the interior product on the left-hand-side is zero. Therefore, $\Psi_1$ is 2-orthogonal to the geminal product.

## Appendix B: Practical implementation of orthogonality constraints

The construction of the configuration space of a GMFCI step is by default as follows: First, a singlet geminal basis set for the active group is generated. At the initial stage, it is a full CI singlet geminal basis set of size $\frac{m(m+1)}{2}$, where $m$ is the number of orbitals, for general guess geminals. However, when singlet-paired, canonical Hartree-Fock orbitals are used as guess geminals, the geminals containing spectator orbitals are removed. Or, if some GMFCI steps have already been performed for the active group considered, the configuration space is spanned by the possibly truncated geminal set obtained in the last step where that group was active.

Suppose without loss of generality that the active group is group 1 and denote this set $(g_1^j)_{j\in\{1,\cdots,M\}}$. Then, one builds the $M \times M$-overlap matrix for the set, $(g_1^j \wedge \Psi_2^0 \wedge \cdots \wedge \Psi_r^0)_{j\in\{1,\cdots,M\}}$, and the set is filtered and orthonormalized. More precisely, the product functions whose squared norm is smaller than the linear dependency tolerance, are filtered out. Next they are Schmidt orthogonalized in turn and their new squared norm is again checked against the linear dependency tolerance. If it is smaller, the configuration is filtered out, if it is larger, the configuration is retained and normalized.

*SF2-orthogonality*

If SF2-orthogonality is imposed, then before the filtering and orthonormalization of the set $(g_1^j \wedge \Psi_2^0 \wedge \cdots \wedge \Psi_r^0)_{j\in\{1,\cdots,M\}}$ as described above, the set $(g_1^j)_{j\in\{1,\cdots,M\}}$ is Schmidt orthogonalized with respect to the normalized set of spectator geminals, that is to say the spectator geminals are projected out orthogonally from the $g_1^j$'s.

*SP2-orthogonality*

If SP2-orthogonality is imposed, then before the filtering and orthonormalization of the set $(g_1^j \wedge \Psi_2^0 \wedge \cdots \wedge \Psi_r^0)_{j\in\{1,\cdots,M\}}$ as described above, the set $(g_1^j)_{j\in\{1,\cdots,M\}}$ is Schmidt orthogonalized with respect to the normalized set of spectator geminals plus natural geminals of each product of pairs of spectator geminals whose population (eigenvalue of the second order reduced density matrix) is above a given threshold (called the "internal geminal threshold"). This is necessary and sufficient according to the lemma of Appendix A to ensure SP2-orthogonality with respect to a geminal product spectator wave function.

Remark 1: Penotti's techniques [23] to impose some orthogonality constraints to sets of orbitals in variational calculations, could be profitably transposed to geminal sets, if in the future a quadratically-convergent algorithm involving energy derivatives is implemented to perform GSCF calculations.

Remark 2: Orthogonality constraints are imposed by restricting the configuration space of the CI calculation. Therefore all the excited states of the active group satisfy the same orthogonality constraints as the optimized ground state.

Remark 3: Note that the lemna of Appendix A is valid irrespective of the spin of the geminals. When dealing with spin-adapted geminals, further simplifications occur. In particular, the orthogonalization with respect to the 2-internal space of a spectator product function, only requires the computation of a basis set singlet 2-internal geminals when the active group is a singlet. This is because the active geminals are already orthogonal by spin symmetry to the triplet 2-internal geminals. Similarly, if the active group is a triplet, only triplet 2-internal geminals need to be extracted from the spectator product. Systems whose total spin is a doublet could be dealt with by adding an extra electron and constraining the active geminal space of the group made of this electron and its finite distance "pseudo companion" to be spanned by functions of the form $g_1^j = \psi_j^\alpha \wedge \psi_\infty^\beta + \psi_\infty^\alpha \wedge \psi_j^\beta$, where $\psi_\infty^\alpha$, $\psi_\infty^\beta$ are fictitious $\alpha-$ and $\beta-$spinorbitals, treated as orthonormal to the other spinorbitals of the basis set $(\psi_j^\alpha, \psi_j^\beta)_j$, and such that the

matrix elements of the Hamiltonian involving any of them are set to zero. This procedure would have the advantages that nothing need to be changed in the implementation of the orthogonality constraints described above, and that when the pseudo pair is spectator, it does not break the spin symmetry of the active group functions.

| | $H_6$ STO-3G | $H_6$ 6-31G | $H_6$ VTZ | $H_8$ STO-3G | $H_{10}$ STO-3G |
|---|---|---|---|---|---|
| ld | $10^{-5}$ | $10^{-5}$ | $10^{-5}$ | $10^{-4}$ | $10^{-4}$ |
| $E^0_{RHF}$ | -3.135532 | -3.227128 | -3.233347 | -4.174370 | -5.214069 |
| $E^0_{step0/CHFOs}$ | -3.159021 | -3.246555 | -3.253214 | -4.195118 | -5.232660 |
| $E^0_{SP1-O}$ | -3.205983 | -3.294840 | -3.302411 | -4.262012 | -5.318586 |
| $E^0_{SP2-O}$ | -3.214018 | -3.301916 | -3.310028 | -4.268768 | -5.325339 |
| SP2-O gem. nb. | 13 | 60 | 144 | 18 | 23 |
| $E^0_{SF2-O}$ | -3.214108 | -3.302039 | -3.310146 | -4.269070 | -5.325612 |
| SF2-O gem. nb. | 17 | 76 | 169 | 30 | 47 |
| $E^0_{GMFCI}$ | -3.214108 | -3.302039 | -3.310146 | -4.269070 | -5.325612 |
| GMFCI gem. nb. | 17 | 78 | 171 | 30 | 47 |
| $E^0_{FullCI}$ | -3.236066 | -3.326551 | -3.335807 | -4.307572 | -5.379955 |

Table 1

GMFCI ground state energies ($E^0$) in Hartree of H-chains with different orthogonality constraints. Adjacent H nuclei are separated by 1 angström. In all calculations, the active group was the one for which the energy lowering was the largest with respect to the guess geminal product function. The threshold on geminal population to decide whether a given natural geminal was considered internal or not, was $10^{-8}$. ld : linear dependency threshold; gem. nb. stands for geminal number, i. e. the size of the CI performed; see main text for the other acronyms and notation.

| | LiH | Be | $Li_2$ | $BeH_2$ | BH | $Be_2$ |
|---|---|---|---|---|---|---|
| $E^0_{RHF}$ | -7.862002 | -14.351880 | -14.638725 | -15.559405 | -24.752780 | -28.698990 |
| $E^0_{step0/CHFOs}$ | -7.882164 | -14.403329 | -14.666525 | -15.570531 | -24.778657 | -28.738672 |
| $E^0_{SF1-O}$ | -7.882203 | -14.403630 | -14.666584 | -15.588630 | -24.807908 | -28.781789 |
| $E^0_{SF2-O}$ | -7.882368 | -14.403654 | -14.667090 | -15.594703 | -24.809920 | -28.803080 |
| $E^0_{GSCF}$ | -7.882372 | -14.403655 | -14.667114 | -15.594715 | -24.809938 | -28.803212 |
| $E^0_{FullCI}$ | -7.882392 | -14.403655 | -14.667340 | -15.594861 | -24.809945 | -28.804345 |

Table 2

GSCF ground state energies ($E^0$) in Hartree at "experimental" geometry ($r_{Li-H} = 1.5957$, $r_{Li-Li} = 2.673$, $r_{Be-H} = 1.340$, $r_{B-H} = 1.2324$, $r_{Be-Be} = 2.460$ in angström) for the STO-3G basis set. $E^0_{SF1-O}$ is the equivalent to $E^0_{SP1-O}$ of Tab. 1. GSCF calculations with ($E^0_{SF2-O}$) and without ($E^0_{GSCF}$) SF2-O constraint were stopped when a GMFCI step did not lower the energy by more than a given threshold, chosen to be $10^{-9}$ Hartree (or $10^{-8}$ for Be and $Be_2$), for all possible active groups. A truncation threshold for quasi-linear dependency of geminal products of $10^{-6}$ has been used, as well as RSSG initial guess, except for LiH where canonical HF geminals were used. See main text for acronyms and notation

| | LiH | BH |
|---|---|---|
| $D^z_{RHF}$ | -4.8578 | 0.9569 |
| $D^z_{step0/CHFOs}$ | -4.6214 | 0.6683 |
| $D^z_{SF1-O}$ | -4.6269 | 0.6861 |
| $D^z_{SF2-O}$ | -4.6189 | 0.6142 |
| $D^z_{GSCF}$ | -4.6197 | 0.6138 |
| $D^z_{FullCI}$ | -4.6201 | 0.6138 |

Table 3

GSCF ground state electric dipole moment along the molecule axis ($D^z$) in Debye for the wave functions corresponding to the energies of Tab. (2).