Error bars and quadratically convergent methods for the numerical simulation of the Hartree-Fock equations

Yvon Maday, Gabriel Turinici

Numerische Mathematik manuscript No.
(will be inserted by the editor)

Error bars and quadratically convergent methods for the numerical simulation of the Hartree-Fock equations

Yvon Maday[1] **and Gabriel Turinici**[2]

[1] Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, 4, Place Jussieu, 75252 Paris Cedex 05, France; e-mail: maday@ann.jussieu.fr
[2] ASCI, UPR 9029, bâtiment 506, Université Paris-Sud, 91405, Orsay Cedex et Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie, France; e-mail: turinici@asci.fr

Received: date / Revised version: date

Summary This paper presents an a posteriori error analysis of the discretization methods used in computational quantum chemistry on the Hartree-Fock equations. Upper and lower bounds for the energy are obtained from any discrete approximation strategy of the solution and the estimator proposed is shown to possess further approximation virtues.

1 Introduction

The purpose of this paper is to present an a posteriori error analysis for the approximation of the Hartree-Fock equations. This analysis is designed to quantitatively asses the performance of an approximation strategy of a solution of the Hartree-Fock equations obtained by prior computation. In agreement with the general paradigm of the a posteriori analysis of [13,15–17], an error bar for an output such as the Hartree-Fock energy starting from the approximated solution at hand is proposed. As in the real laboratory experiments, numerical computations do not provide the exact value of the searched quantity but rather an approximation that is to be qualified by the use of the error bars in exactly the same spirit as in the laboratory measurements. In addition we will show that in some cases the a posteriori method may also be seen as an accelerator of the convergence of the primary algorithm used to compute the solution.

The time independent Schrödinger equation that models the behavior of a quantum molecular system deals with state functions $\psi(\underline{x})$, where $\underline{x}$ denotes the position of the particles (nuclei and electrons) hence is a variable that lives in $\mathbb{R}^{3K}$ where K is the number of particles[1]. This system is far too large to be directly tractable by numerical simulations for molecules larger than the hydrogen atom. The quantum chemists have thus introduced a series of simplified models. One of them (the Born Oppenheimer approximation) allows to separate the electron and the nuclei so as to consider first a system in which only the N electrons of the molecule move (thus are the only N variables of the state function) and the nuclei are fixed in $\overline{x}_j$ (and appear as parameters). For each configuration $(\overline{x}_1, ..., \overline{x}_m)$ of the m nuclei a complex electronic wavefunction $\Phi(x_1, ..., x_N) \in \mathbb{C}$, $x_i \in \mathbb{R}^3$, $i = 1, ..., N$ is sought after that minimizes the energy of the system. This first simplification is nevertheless not sufficient to make the resulting equations accessible for computations for large molecules; another simplification is therefore introduced by considering that the state function is a N dimensional determinant of simple functions of $\mathbb{R}^3$, called *Slater determinant*:

$$\Phi(r_1, ..., r_N) = \frac{1}{\sqrt{N!}} \det(\Phi_i(r_j)),$$

where Φ_i, $i = 1, ..., N$ are now functions of one variable in $\mathbb{R}^3$ chosen orthogonal with respect to the canonical scalar product $< \cdot, \cdot >$ on $L^2(\mathbb{R}^3)$.

Let us denote by $\mathcal{K}$ the subset of $(L^2(\mathbb{R}^2))^N$ defined by

$$\mathcal{K} = \{(\Phi_1, ..., \Phi_N) \in (L^2(\mathbb{R}^2))^N; < \Phi_i, \Phi_j >= \delta_{ij}\}. \tag{1}$$

Assuming that the molecule is isolated and only Coulombic forces are present, the description of the non-relativistic electrons where, for the sake of simplicity we have neglected the spin dependency, leads to the following expression of the Hartree-Fock energy :

$$\mathcal{E}^{HF}(\Phi_1, ..., \Phi_N) = \sum_{i=1}^{N} \int_{\mathbb{R}^3} \left(|\nabla \Phi_i|^2 + V \ |\Phi_i|^2 \right) +$$
$$\frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\Phi(x) \rho_\Phi(y)}{|x - y|} dxdy - \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|\rho_\Phi(x, y)|^2}{|x - y|} dxdy, \tag{2}$$

[1] we will consider non relativistic models without spin variables

where the *density matrix* $\rho_\Phi(x,y)$, the *electronic density* $\rho_\Phi(x)$ and the *potential* V are given by the formulaes :

$$\begin{aligned}\rho_\Phi(x,y) &= \sum_{i=1}^{N} \Phi_i(x)\overline{\Phi_i(y)} \\ \rho_\Phi(x) &= \rho_\Phi(x,x) \\ V(x) &= -\sum_{j=1}^{m} \frac{Z_j}{|x-\overline{x}_j|}.\end{aligned} \tag{3}$$

We have denoted here by $Z_j > 0$ the charge of the j-th nucleo.

In order to determine the ground state of the molecule that, by definition, minimizes the energy (2) under the constraint (1), the Euler Lagrange equations give rise to the Hartree-Fock problem :

Find a $L^2(\mathbb{R}^3)$-orthonormal system $\Phi = \{\Phi_i\}_{i=1,N}^t$ and an hermitian matrix $\Lambda = [\lambda_{i,j}]_{i,j=1,N}$ such that

$$\forall i,\ 1 \le i \le N, \quad \mathcal{F}_\Phi(\Phi_i) = -\sum_{j=1}^{N} \lambda_{i,j}\Phi_j, \tag{4}$$

where $\mathcal{F}_\Phi$ is the Fock operator. When acting on an element ψ regular enough of the variable $x \in \mathbb{R}^3$, this operator associates the following function of the $x \in \mathbb{R}^3$ variable:

$$\mathcal{F}_\Phi(\psi)(x) = \Big(-\Delta + V(x) + (\rho_\Phi \star \frac{1}{|x|})\Big)\psi(x) - \int_{\mathbb{R}^3} \frac{\rho_\Phi(x,y)}{|x-y|}\psi(y)dy. \tag{5}$$

Here $\star$ is the convolution product

$$(f \star g)(x) = \int_{\mathbb{R}^3} f(x-y)g(y)dy.$$

Remark 1 It is standard to notice that the density matrix is invariant under unitary transforms, i.e. for any element U of the set of the $N \times N$ unitary matrices $\mathcal{U}(N)$:

$$\forall (x,y) \in \mathbb{R}^3, \quad \rho_\Phi(x,y) = \rho_{U\Phi}(x,y) \tag{6}$$

Hence it follows that the unitary transform U can be chosen in such a way that the hermitian matrix Λ become diagonal: $\Lambda = [\lambda_i]_{i=1,N}$. The solution $\Psi = U\Phi = \{(U\Phi)_i\}_{i=1,N}$ satisfies indeed the more simple Hartree-Fock problem :

$$\forall i,\ 1 \le i \le N, \quad \mathcal{F}_\Psi(\psi_i) = -\lambda_i\psi_i \tag{7}$$

The problem then appears as a non linear eigenvalue problem.

This highly nonlinear problem is solved through iterations known as Self Consistent Field approximation; we refer to [6] for a very recent and complete analysis on the convergence of some of these algorithms (Roothaan algorithm and the level shifting algorithm). It is still a very expensive problem since the non linear contribution has a large computational complexity (we refer to [20, 8] for some example of tailored techniques to minimize this complexity). The numerical analysis of the method used typically by the chemists community is most often an open problem and in any case will not provide sound information since most of the numerical approximations are very often at the limit of the convergence. More interesting seems the concept of *a posteriori error estimators* where, from the computed solution, it is possible to derive reliable information about the validity of the computation that has been done. The purpose of this paper is in this direction.

Denote by $\mathcal{H} = (H^1(\mathbb{R}^3))^N$ the natural space for the solutions of the Hartree-Fock equations and by F_{ij} the mapping $F_{ij} : \mathcal{H} \mapsto \mathbb{R}$ defined over any element $\Phi = (\Phi_i)_{i=1}^N$ by

$$F_{ij}(\Phi) = < \Phi_i, \Phi_j > -\delta_{ij}.$$

In all that follows any N-tuple element $\Phi = (\Phi_i)_{i=1}^N$ will be supposed to be a **column** $(N \times 1)$ vector of $\mathcal{H}$. Consider the minimization problem

$$\inf\{\mathcal{E}^{HF}(\Phi); \Phi \in \mathcal{H} \cap \mathcal{K}\} \tag{8}$$

Remark 2 The analysis of problem (7) is not completely under control: we can cite the partial results obtained in [10, 11] about the existence of a ground state for positive or neutral molecules and non existence results for negative ions. The basic result of uniqueness of the density solution is still an open problem of outstanding difficulty. Under the hypothesis

$$\sum_{j=1}^{m} Z_j > N - 1, \tag{9}$$

it has been proven in [11] that a minimum of the problem (8) exists and any such minimum is a solution of the Hartree-Fock equation (4). Moreover, when this problem is written in the form (7) additional information is available on λ_i, namely $\lambda_i > 0$, $i = 1, ..., N$. We will assume in all that follows that (9) is true.

In order to make the presentation easy, we will assume in all that follows that the electronic wavefunction is real and will work on real function spaces; trivial adaptations allow the treatment of complex valued wavefunctions.

2 Error decomposition

2.1 Error metrics

Let $\Phi_0 = (\Phi_{0i})_{i=1}^N \in \mathcal{H} \cap \mathcal{K}$ be a minimum of (8) and $\Phi = (\Phi_i)_{i=1}^N \in \mathcal{H} \cap \mathcal{K}$ an approximation of Φ_0 obtained as the solution of a minimization problem:

$$\inf\{\mathcal{E}^{HF}(\Phi); \Phi \in X^N \cap \mathcal{K}\} \tag{10}$$

where X is a finite dimensional subspace of $H^1(\mathbb{R}^3)$.

The a posteriori analysis on the one hand studies bounds for the difference $\mathcal{E}^{HF}(\Phi_0) - \mathcal{E}^{HF}(\Phi)$ and on the other hand proposes explicit trust intervals on the desired (but unknown) quantity $\mathcal{E}^{HF}(\Phi_0)$ using only the approximate solution at hand Φ ; of course, due to the variational setting, an upper bound on $\mathcal{E}^{HF}(\Phi_0)$ is $\mathcal{E}^{HF}(\Phi)$ itself; the main focus will therefore be placed on finding lower bounds for $\mathcal{E}^{HF}(\Phi_0)$, which is a non-trivial problem that, to our knowledge, has not been addressed in the literature.

Before dwelling into the a posteriori analysis of (8) it is crucial to introduce the proper definition for the error between a minimizer Φ_0 and its approximation Φ. To this end one has to recall the invariance property of the Hartree-Fock energy:

$$\mathcal{E}^{HF}(\Psi) = \mathcal{E}^{HF}(U\Psi), \forall \Psi \in \mathcal{H} \cap \mathcal{K},\ \forall U \in \mathcal{U}(N) \tag{11}$$

From (11) it follows that if Φ_0 is a minimizer of (8), then for any $U \in \mathcal{U}(N)$, $U\Phi_0$ is also a minimizer and therefore a solution of (4). The same considerations remain true for the problem (10). It is therefore natural to consider the distance between the sets $\{U\Phi_0; U \in \mathcal{U}(N)\}$ and $\{V\Phi; V \in \mathcal{U}(N)\}$ as the most appropriate definition of the distance between Φ_0 and Φ. For reasons that will be made clear later on, we will use in fact an equivalent form (see section 2.3) of the above definition. For any $\Psi_1, \Psi_2 \in \mathcal{H}$ let

$$U_{\Psi_1,\Psi_2} = argmin\{\|U\Psi_1 - \Psi_2\|^2_{(L^2(\mathbb{R}^3))^N}; U \in \mathcal{U}(N)\}. \tag{12}$$

For a given norm $\|\cdot\|$ ($\|\cdot\|_{(L^2)^N}$, $\|\cdot\|_{(H^1)^N}$ ...) we will measure the distance between (sets represented by) Ψ_1 and Ψ_2 as:

$$\|\Psi_1 - \Psi_2\|_\star = \|U_{\Psi_1,\Psi_2}\Psi_1 - \Psi_2\| = \|\Psi_1 - U_{\Psi_2,\Psi_1}\Psi_2\|,$$

the last equality being motivated by the fact that $U_{\Psi_2,\Psi_1} = U^t_{\Psi_1,\Psi_2} \in \mathcal{U}(N)$.

Remark 3 Note from (12) that U_{Ψ_2,Ψ_1} is intrinsically related to the norm of $(L^2)^N$; when $\|\cdot\| = \|\cdot\|_{(L^2)^N}$ we recover the distance between the sets $\{U\Psi_1; U \in \mathcal{U}(N)\}$ and $\{V\Psi_2; V \in \mathcal{U}(N)\}$.

The properties of this metric are closely related to the following decomposition of $\mathcal{H}$:

$$\mathcal{H} = \mathcal{A}_\Phi \oplus \mathcal{S}_\Phi \oplus \Phi^{\perp\!\!\!\perp} \tag{13}$$

where for any $\Phi \in \mathcal{H} \cap \mathcal{K}$:

$$\begin{aligned}
\mathcal{A}_\Phi &= \{C\Phi; C \in \mathbb{R}^{N\times N}, C^t = -C\} \\
\mathcal{S}_\Phi &= \{S\Phi; S \in \mathbb{R}^{N\times N}, S^t = S\} \\
\Phi^{\perp\!\!\!\perp} &= \{\Psi = (\Psi_i)_{i=1}^N \in \mathcal{H}; < \Psi_i, \Phi_j >= 0; i,j = 1,...,N\}
\end{aligned}$$

We will denote for any $\Psi_1, \Psi_2 \in (L^2)^N$: $\Psi_1 \perp\!\!\!\perp \Psi_2$ if for any $i,j = 1,N$: $< (\Psi_1)_i, (\Psi_2)_j >= 0$; then $\Phi^{\perp\!\!\!\perp}$ can be defined equivalently

$$\Phi^{\perp\!\!\!\perp} = \{\Psi \in \mathcal{H}; \Psi \perp\!\!\!\perp \Phi\}.$$

For any $\xi = (\xi_i)_{i=1}^N \in \mathcal{H}$ the decomposition (13) is obtained in the following manner: compute the matrix $M = (M_{ij})_{i,j=1}^N$ where for each $i,j = 1,...,N$: $M_{ij} =< \xi_i, \Phi_j >$. Denote by S the symmetric part of M: $S = \frac{M+M^t}{2}$ and by C the antisymmetric part: $C = \frac{M-M^t}{2}$. Then $S\Phi$ will be the component of ξ in the space $\mathcal{S}_\Phi$ and $C\Phi$ the component of ξ in the space $\mathcal{A}_\Phi$; in addition it is easy to see that $(\xi - S\Phi - C\Phi) \perp\!\!\!\perp \Phi$, so the difference $\xi - S\Phi - C\Phi$ is in $\Phi^{\perp\!\!\!\perp}$.

Lemma 1 *Let $\Phi, \Psi \in \mathcal{H} \cap \mathcal{K}$. Then the matrix $U_{\Psi,\Phi}$ solution of (12) has the properties*

$$U_{\Psi,\Phi}\Psi - \Phi \in \mathcal{S}_\Phi \oplus \Phi^{\perp\!\!\!\perp}, \quad \Phi - U_{\Psi,\Phi}\Psi \in \mathcal{S}_{U_{\Psi,\Phi}\Psi} \oplus \Psi^{\perp\!\!\!\perp}. \tag{14}$$

In particular for $\Psi = \Phi_0$,

$$U_{\Phi_0,\Phi}\Phi_0 = \Phi + S\Phi + W, \ S \in \mathbb{R}^{N\times N} : S^t = S, \ W \in \Phi^{\perp\!\!\!\perp}. \tag{15}$$

Proof. Consider the decomposition

$$\Psi - \Phi = C\Phi + S\Phi + W, \ C\Phi \in \mathcal{A}_\Phi, \ S\Phi \in \mathcal{S}_\Phi, \ W \in \Phi^{\perp\!\!\!\perp}, \tag{16}$$

and denote $M = C + S$. Then we can write

$$\begin{aligned}
&U_{\Psi,\Phi} = argmin\{\|U\Psi - \Phi\|^2_{(L^2(\mathbb{R}^3))^N}; U \in \mathcal{U}(N)\} \\
&= argmin\{\|U((Id_N + M)\Phi + W) - \Phi\|^2_{(L^2(\mathbb{R}^3))^N}; U \in \mathcal{U}(N)\} \\
&= argmin\{\|(U(Id_N + M) - Id_N)\Phi\|^2_{(L^2(\mathbb{R}^3))^N}; U \in \mathcal{U}(N)\} \\
&= argmin\{\|U(Id_N + M) - Id_N)\|^2_{\mathbb{R}^{N\times N}}; U \in \mathcal{U}(N)\} \\
&= argmin\{\|(Id_N + M) - U^t\|^2_{\mathbb{R}^{N\times N}}; U \in \mathcal{U}(N)\} \qquad (17)
\end{aligned}$$

The transformation from the second to the third line is a consequence of the fact that $W \perp\!\!\!\perp \Phi$ so therefore $U(Id_N + M)\Phi \perp\!\!\!\perp W$; the next equality is true because $\Phi \in \mathcal{K}$.
For any antisymmetric matrix $\tilde{C} \in \mathbb{R}^{N\times N}$ consider the path in $\mathcal{U}(N)$ given by $t \to e^{\tilde{C}t}U_{\Psi,\Phi}$. The tangent at $t = 0$ to this path is $\tilde{C}U_{\Psi,\Phi}$. Writing the first order conditions for the minimality in (17) we obtain:

$$\begin{aligned}
0 =& < (Id_N + M) - U^t_{\Psi,\Phi}, U^t_{\Psi,\Phi}\tilde{C}^t >_{\mathbb{R}^{N\times N}} \\
=& < U_{\Psi,\Phi}(Id_N + M) - Id_N, \tilde{C}^t >_{\mathbb{R}^{N\times N}}, \\
& \forall \tilde{C} \in \mathbb{R}^{N\times N} : \tilde{C}^t = -\tilde{C},
\end{aligned}$$

which shows that $U_{\Psi,\Phi}(Id_N + M)$ is a symmetric matrix ; and therefore $U_{\Psi,\Phi}\Psi \in \mathcal{S}_\Phi \oplus \Phi^{\perp\!\!\!\perp}$. To prove the second part of the equation (14) denote for any Ψ_1, Ψ_2 by C_{Ψ_1,Ψ_2} the antisymmetric matrix appearing in the decomposition $\Psi_1 - \Psi_2 = C_{\Psi_1,\Psi_2}\Psi_2 + S_{\Psi_1,\Psi_2}\Psi_2 + W_{\Psi_1,\Psi_2}$ with $C_{\Psi_1,\Psi_2}\Psi_2 \in \mathcal{A}\Psi_2$, $S_{\Psi_1,\Psi_2}\Psi_2 \in \mathcal{S}\Psi_2$ and $W_{\Psi_1,\Psi_2} \in \Psi_2^{\perp\!\!\!\perp}$; then one obtains by straightforward computations $C_{\Psi_1,\Psi_2} = -C_{\Psi_2,\Psi_1}$. □

Remark 4 In practice the representative of the class of isoenergy functions $\{U\Phi_0; U \in \mathcal{U}(N)\}$ is taken to be the one that solves equations (7), and the same is true for any of its approximations Φ. It is not clear whether a norm for which this practical choice gives optimal approximations in the sense of (12) exists and to what extent this choice is also optimal in the L^2 norm.

2.2 Order of the symmetric part of the error

Let $\Psi, \Phi \in \mathcal{H} \cap \mathcal{K}$ and let us consider the decomposition (16). We have seen that the antisymmetric part given by matrix C may be set to zero modulo some appropriate "rotation" on Ψ ; it is therefore natural to study the properties of the symmetric part $S\Phi$.

Lemma 2 *Let $\Psi, \Phi \in \mathcal{H} \cap \mathcal{K}$ with associated decomposition (16). Then there exists constants C_1, C_2 depending only of N such that:*

$$\|S\Phi\|_{(L^2(\mathbb{R}^3))^N} \leq C_1 \|\Psi - \Phi\|^2_{(L^2(\mathbb{R}^3))^N} \tag{18}$$

$$\|S\Phi\|_{\mathcal{H}} \leq C_2 \|\Psi - \Phi\|^2_{\mathcal{H}} \|\Phi\|_{\mathcal{H}} \tag{19}$$

Proof. Let us write $W = D\tilde{W}$ such that $< \tilde{W}_i, \tilde{W}_j >= \delta_{ij}$, $M = C + S$. Denote

$$\epsilon = \|\Psi - \Phi\|_{(L^2(\mathbb{R}^3))^N} = \sqrt{\sum_{i,j=1}^{N} M_{ij}^2 + D_{ij}^2}$$

Since $\Psi \in \mathcal{K}$, $F_{ij}(\Psi) = 0$, $i, j = 1, ..., N$. For $j = i$ we obtain:

$$1 = (1 + M_{ii})^2 + \sum_{j \neq i} M_{ij}^2 + \sum_{j=1}^{N} D_{ij}^2,$$

or equivalently:

$$S_{ii} = M_{ii} = -\frac{\sum_{j=1}^{N} M_{ij}^2 + \sum_{j=1}^{N} D_{ij}^2}{2},$$

which proves that $M_{ii} \leq \epsilon^2$, $i = 1, ..., N$. For $i \neq j$ one obtains:

$$0 = \sum_{k \neq i, k \neq j} M_{ik} M_{jk} + (M_{ii} + 1) M_{ji} + M_{ij}(M_{jj} + 1) + \sum_{k=1}^{N} D_{ki} D_{kj},$$

which gives after straightforward manipulations $S_{ij} = \frac{M_{ij} + M_{ji}}{2} \leq \epsilon^2$; this concludes the proof of (18). For (19) one denotes first that $\|\Psi - \Phi\|_{(L^2(\mathbb{R}^3))^N} \leq \|\Psi - \Phi\|_{\mathcal{H}}$ and apply (18) to conclude that $S_{ij} \leq \|\Psi - \Phi\|^2_{\mathcal{H}}$, $i, j = 1, ..., N$. The conclusion follows then by the definition of the norm $\|\cdot\|_{\mathcal{H}}$. □

2.3 Optimality in H^1 norm

We have proposed in section 2.1 that for any norm $\|\cdot\|$ the error $\Phi_0 - \Phi$ be computed as $\|U_{\Phi_0,\Phi}\Phi_0 - \Phi\|$. Since the definition $U_{\Phi_0,\Phi}$ is closely related to the L^2 norm it is natural to ask whether this definition is still appropriate when norms other than L^2 are used, for instance the canonical norm of $\mathcal{H}$. The situation is settled by the following

Lemma 3 *Let $\Psi = (\Psi_1, ..., \Psi_N) \in \mathcal{H} \cap \mathcal{K}$ and $\Phi \in \mathcal{H} \cap \mathcal{K}$ and denote*

$$U^1_{\Psi,\Phi} = argmin\{\|U\Psi - \Phi\|_{\mathcal{H}}; U \in \mathcal{U}(N)\}$$

There exists a constant c depending only of N and Ψ such that

$$c\|U_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}} \leq \|U^1_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}} \leq \|U_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}}$$

Proof. The inequality

$$\|U_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}} \geq \|U^1_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}}$$

follows as a consequence of the definition of $U^1_{\Psi,\Phi}$.

Denote by F the linear space generated by $\{\Psi_1, ..., \Psi_N\}$ and define:

$$M = \{\zeta \in H^1(\mathbb{R}^3); < \zeta, \chi >_{L^2,L^2} = 0, \ \forall \chi \in F\}.$$

For any $\chi \in H^1(\mathbb{R}^3)$ denote by χ_F the L^2 projection of χ on F and $\chi_M = \chi - \chi_F$. We define a norm $\|\cdot\|_d$ on $H^1(\mathbb{R}^3)$ as follows:

$$\|\chi\|^2_d = \|\chi_F\|^2_{L^2} + \|\chi_M\|^2_{H^1(\mathbb{R}^3)}.$$

We will prove that this norm is equivalent to the canonical norm of $H^1(\mathbb{R}^3)$ (with constants depending only on N and Ψ). Write for any $\chi \in H^1(\mathbb{R}^3)$:

$$\begin{aligned}\|\chi\|_{H^1(\mathbb{R}^3)} &\leq \|\chi - \chi_F\|_{H^1(\mathbb{R}^3)} + \|\chi_F\|_{H^1(\mathbb{R}^3)} \\ &\leq \|\chi\|_d + \|\chi_F\|_{H^1(\mathbb{R}^3)} \leq C\|\chi\|_d\end{aligned}$$

where we have used the fact that the norms $\|\cdot\|_{L^2}$ and $\|\cdot\|_{H^1(\mathbb{R}^3)}$ are equivalent on the **finite dimensional** space F. It follows that there exists a constant C (depending only on N and Ψ) such that for any $\chi \in H^1(\mathbb{R}^3)$

$$\|\chi\|_{H^1(\mathbb{R}^3)} \leq C\|\chi\|_d.$$

We will prove next that the norm $\|\cdot\|_{H^1(\mathbb{R}^3)}$ can also be lower bounded by the norm $\|\cdot\|_d$ modulo some constant depending only N and Ψ. Assume on the contrary that this is not true. Then there exists a sequence $(\chi_n)_{n\geq 1} \subset H^1(\mathbb{R}^3)$ such that $\|\chi_n\|_d = 1$ and $\|\chi_n\|_{H^1(\mathbb{R}^3)} \to 0$ as $n \to \infty$. It follows that the sequence χ_n converges to zero in L^2 and in particular the sequence $(\chi_{nF})_{n\geq 1}$ of L^2 projections to F is also converging to zero: $\|\chi_{nF}\|_{L^2} \to 0$ $(n \to \infty)$; by the same argument as above we obtain $\|\chi_{nF}\|_{H^1(\mathbb{R}^3)} \to 0$ $(n \to \infty)$. Then

$$\|\chi_{nM}\|_{H^1(\mathbb{R}^3)} = \|\chi_n - \chi_{nF}\|_{H^1(\mathbb{R}^3)} \leq \|\chi_n\|_{H^1(\mathbb{R}^3)} + \|\chi_{nF}\|_{H^1(\mathbb{R}^3)}$$

and it follows that $\|\chi_{nM}\|_{H^1(\mathbb{R}^3)} \to 0$ $(n \to \infty)$. Together with $\|\chi_{nF}\|_{L^2} \to 0$ $(n \to \infty)$ we conclude that $\|\chi_n\|_d \to 0$ $(n \to \infty)$, in

contradiction with the initial assumption. We have therefore proved that there exists constants c, C (depending only N and Ψ) such that for any $\chi \in H^1(\mathbb{R}^3)$

$$c\|\chi\|_d \leq \|\chi\|_{H^1(\mathbb{R}^3)} \leq C\|\chi\|_d.$$

The above equivalence imply that the canonical norm $\|\cdot\|_{d,N}$ of $(H^1(\mathbb{R}^3), \|\cdot\|_d)^N$ is equivalent (with constants depending only on N and Ψ) to the canonical norm of $\mathcal{H}$:

$$c_1\|\zeta\|_{\mathcal{H}} \leq \|\zeta\|_{d,N} \leq C_1\|\zeta\|_{\mathcal{H}},\ \forall \zeta \in \mathcal{H}.$$

Since $\Psi \in \mathcal{K}$, the functions $\{\Psi_1, ..., \Psi_N\}$ are orthonormal with respect to the scalar product of $L^2(\mathbb{R}^3)$ and also with respect to the scalar product $<\cdot,\cdot>_d$ associated with the norm $\|\cdot\|_d$. It follows by (17) that

$$U_{\Psi,\Phi} = argmin\{\|U\Psi - \Phi\|_{\|\cdot\|_{d,N}}; U \in \mathcal{U}(N)\},$$

as both solve the same minimization problem on $\mathcal{U}(N)$. But then

$$\|U^1_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}} \geq \frac{1}{C_1}\|U^1_{\Psi,\Phi}\Psi - \Phi\|_{d,N} \geq$$
$$\frac{1}{C_1}\|U_{\Psi,\Phi}\Psi - \Phi\|_{d,N} \geq \frac{c_1}{C_1}\|U_{\Psi,\Phi}\Psi - \Phi\|_{\mathcal{H}}.$$

which concludes the proof. □

3 Optimality conditions and coercivity

We will begin this section with some elementary information about the geometry of the manifolds $\mathcal{K}$ and $\mathcal{H} \cap \mathcal{K}$:

Lemma 4 *Let $\Phi \in \mathcal{H} \cap \mathcal{K}$. The tangent space in Φ to the manifold $\mathcal{H} \cap \mathcal{K}$ is $\mathcal{A}_\Phi \oplus \Phi^{\perp\!\!\!\perp}$.*

Proof. Let $\Phi(t) :]-\epsilon, \epsilon[\to \mathcal{H} \cap \mathcal{K}$, $\epsilon > 0$, $\Phi(0) = \Phi$ be a C^1 path in $\mathcal{H} \cap \mathcal{K}$. Consider the decomposition $\Phi'(0) = S\Phi + C\Phi + W$, $S\Phi \in \mathcal{S}_\Phi$, $C\Phi \in \mathcal{A}_\Phi$, $W \in \Phi^{\perp\!\!\!\perp}$. By differentiating the condition $F_{ij}(\Phi(t)) = 0$ we obtain $< \Phi_i, \Phi'_j(0) > + < \Phi'_i(0), \Phi_j >= 0$ which proves that $S_{ij} = 0$. Since this is true for any $i, j = 1, ..., N$ we conclude $S = 0$ i.e. $\Phi'(0) \in \mathcal{A}_\Phi \oplus \Phi^{\perp\!\!\!\perp}$.

To prove that any $\Psi = C\Phi + W \in \mathcal{A}_\Phi \oplus \Phi^{\perp\!\!\!\perp}$ may be seen as the tangent in Φ of a C^1 path in $\mathcal{H} \cap \mathcal{K}$, choose $\Phi(t) :]-\epsilon, \epsilon[\to \mathcal{H} \cap \mathcal{K}$, $0 < \epsilon < 1$, $\Phi(t) = \sqrt{1-t^2}e^{Ct}\Phi + tW$ and note that $\Phi'(0) = \Psi$ and $\|\Phi(t)\| = 1,\ -\epsilon < t < \epsilon$. □

The second order optimality conditions for the minimization problem (8) will be seen to be very useful within our approach. Let $\Phi_0 \in \mathcal{H} \cap \mathcal{K}$ be a minimizer of (8) and Λ^0 be the hermitian matrix corresponding to Φ_0 in equations (4). We will write the second order conditions in the form:

$$D^2\mathcal{E}^{HF}(\Phi_0)(\Psi,\Psi)+ < \Lambda^0\Psi,\Psi >_{(L^2(\mathbb{R}^3))^N} \geq 0,\ \forall \Psi \in \mathcal{A}_{\Phi_0} \oplus \Phi_0^{\perp\!\!\!\perp}.$$

Denote for any $\Phi \in \mathcal{H} \cap \mathcal{K}$:

$$\mathcal{E}^{\Phi}(\cdot) = \mathcal{E}^{HF}(\cdot) + \sum_{i,j=1}^{N} \Lambda_{ij} F_{ij}(\cdot) \tag{20}$$

where $\Lambda_{ij} =< \mathcal{F}_{\Phi}\Phi_i, \Phi_j >$, $i,j = 1,...,N$.

Remark 5 The Hartree-Fock equations (4) can be "symbolically" derived as a corollary of lemma 4. Indeed, the first order minimality conditions associated to (8) read

$$< D\mathcal{E}^{HF}(\Phi_0), \Psi >_{(L^2(\mathbb{R}^3))^N} = 0,\ \forall \Psi \in \mathcal{A}_{\Phi_0} \oplus \Phi_0^{\perp\!\!\!\perp}$$

which is the same as writing $D\mathcal{E}^{HF}(\Phi_0) = S\Phi_0$, ($S$ being a symmetric matrix) which are exactly equations (4) since $D\mathcal{E}^{HF}(\Phi_0)$ can be identified with $(\mathcal{F}_{\Phi_0},...,\mathcal{F}_{\Phi_0})$. Moreover, with the definition (20) we note that

$$D\mathcal{E}^{\Phi_0} \equiv 0. \tag{21}$$

Denote by $a_{\Phi}(\cdot,\cdot)$ the bilinear form $D^2\mathcal{E}^{\Phi}(\Phi)(\cdot,\cdot)$ and remark that

$$a_{\Phi_0}(\cdot,\cdot) = D^2\mathcal{E}^{HF}(\Phi_0)(\cdot,\cdot)+ < \Lambda^0\cdot,\cdot >_{(L^2(\mathbb{R}^3))^N}\ .$$

In order to obtain an explicit formula for a_{Φ_0} we need the expression of $D^2\mathcal{E}^{HF}(\Phi_0)$. Let $\Phi, \Psi^1, \Psi^2 \in \mathcal{H} \cap \mathcal{K}$. Then

$$\begin{aligned}
D^2\mathcal{E}^{HF}(\Phi)(\Psi^1,\Psi^2) &= 2\cdot\sum_{i=1}^{N}\int_{\mathbb{R}^3}\left(\nabla\Psi_i^1\cdot\nabla\Psi_i^2 + V\Psi_i^1\Psi_i^2\right)\\
&+\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{8\rho_{\Phi,\Psi^1}(x)\rho_{\Phi,\Psi^2}(y)+4\rho_{\Psi^1,\Psi^2}(x)\rho_{\Phi}(y)}{|x-y|}dxdy\\
&-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{2\rho_{\Phi}(x,y)(\rho_{\Psi^1,\Psi^2}(x,y)+\rho_{\Psi^1,\Psi^2}(y,x))}{|x-y|}dxdy\\
&-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{+4\rho_{\Phi,\Psi^1}(x,y)(\rho_{\Phi,\Psi^2}(x,y)+\rho_{\Phi,\Psi^2}(y,x))}{|x-y|}dxdy,
\end{aligned}$$

with the definitions

$$\rho_{\Psi^1,\Psi^2}(x,y) = \sum_{i=1}^N \Psi_i^1(x)\Psi_i^2(y),$$
$$\rho_{\Psi^1,\Psi^2}(x) = \rho_{\Psi^1,\Psi^2}(x,x).$$

We obtain therefore:

$$\begin{aligned}
D^2\mathcal{E}^{HF}(\Phi_0)(\Psi,\Psi) &= 2\cdot\sum_{i=1}^N \int_{\mathbb{R}^3}\left(|\nabla\Psi_i|^2 + V\Psi_i^2\right)\\
&+\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{8\rho_{\Phi_0,\Psi}(x)\rho_{\Phi_0,\Psi}(y)+4\rho_{\Psi}(x)\rho_{\Phi_0}(y)}{|x-y|}dxdy\\
&-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{4\rho_{\Phi_0}(x,y)\rho_{\Psi}(x,y)}{|x-y|}dxdy\\
&-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{4\rho_{\Phi_0,\Psi}(x,y)(\rho_{\Phi_0,\Psi}(x,y)+\rho_{\Phi_0,\Psi}(y,x))}{|x-y|}dxdy.
\end{aligned}$$

We will study in the following the coercivity properties of the bilinear form a_{Φ_0}. Note that for any $\Psi\in\mathcal{H}\cap\mathcal{K}$: $\mathcal{E}^{HF}(\Psi)=\mathcal{E}^{\Phi_0}(\Psi)$ and in addition $a_{\Phi_0}=D^2\mathcal{E}^{\Phi_0}(\Phi_0)$. By differentiating the invariance property (11) we obtain in particular (cf. lemma 4):

$$D\mathcal{E}^{\Phi_0}(\Psi)(C\Psi)=0, \forall\Psi\in\mathcal{H}\cap\mathcal{K},\ \forall C\Psi\in\mathcal{A}_\Psi. \tag{22}$$

Differentiating now (22) in $\Psi=\Phi_0$ and taking into account the fact that Φ_0 is a solution of (4) we obtain:

$$< D^2\mathcal{E}^{\Phi_0}(\Phi_0)(C\Phi_0,\tilde{C}\Phi_0+W)=0,\ \forall C\Phi_0,\tilde{C}\Phi_0\in\mathcal{A}_{\Phi_0},\ \forall W\in\Phi_0^{\perp\!\!\!\perp}.$$

Then it follows that a_{Φ_0} vanishes on $\mathcal{A}_{\Phi_0}$ thus cannot be coercive there ; the coercivity properties of a_{Φ_0} are described by the following two lemmata.

Lemma 5 *Let V_{Φ_0} be the closure of span$\{\Psi\in\mathcal{A}_{\Phi_0}\oplus\Phi_0^{\perp\!\!\!\perp} : a_{\Phi_0}(\Psi,\Psi)=0\}$ with respect to the canonical topology of $\mathcal{H}$. Then a_{Φ_0} is null on $V_{\Phi_0}\times V_{\Phi_0}$.*

Proof. Let $\Psi^1,\Psi^2\in\mathcal{A}_{\Phi_0}\oplus\Phi_0^{\perp\!\!\!\perp}$ be such that $a_{\Phi_0}(\Psi^i,\Psi^i)=0$, $i=1,2$. Then since $a_{\Phi_0}\geq 0$ on $\mathcal{A}_{\Phi_0}\oplus\Phi_0^{\perp\!\!\!\perp}$ by a standard Cauchy-Schwartz inequality for the positive bilinear form a_{Φ_0} we obtain $2|a_{\Phi_0}(\Psi^1,\Psi^2)|\leq a_{\Phi_0}(\Psi^1,\Psi^1)+a_{\Phi_0}(\Psi^2,\Psi^2)$ and therefore $a_{\Phi_0}(\Psi^1,\Psi^2)=0$. It follows then that for any $\Psi=\mu_1\Psi^1+\mu_2\Psi^2$ such that $\mu_1,\mu_2\in\mathbb{R}$ we have $a_{\Phi_0}(\Psi,\Psi)=0$ which, together with the continuity of a_{Φ_0} concludes the proof. □

Proposition 1 *Let X_{Φ_0} be a closed subspace of $\Phi_0^{\perp\!\!\!\perp}$ $(\mathcal{H})$ such that*

$$\forall \Psi \in X_{\Phi_0}, \Psi \neq 0 : \ a_{\Phi_0}(\Psi, \Psi) > 0.$$

Then a_{Φ_0} is coercive on X_{Φ_0}.

The proof of this proposition makes use of the following auxiliary result

Lemma 6 *The mapping*

$$\Psi \mapsto \frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{8\rho_{\Phi_0,\Psi}(x)\rho_{\Phi_0,\Psi}(y) + 4\rho_\Psi(x)\rho_{\Phi_0}(y)}{|x-y|}dxdy$$
$$-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{4\rho_{\Phi_0}(x,y)\rho_\Psi(x,y)}{|x-y|}dxdy$$
$$-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{4\rho_{\Phi_0,\Psi}(x,y)(\rho_{\Phi_0,\Psi}(x,y) + \rho_{\Phi_0,\Psi}(y,x))}{|x-y|}dxdy$$

is sequentially weakly lower semicontinuous with respect to the canonic topology of $\mathcal{H}$.

Proof. Let us recall the Hardy inequality (used in the version of [11] p.42) which holds for all $y \in \mathbb{R}^3$, $\varphi \in H^1(\mathbb{R}^3)$:

$$\int_{\mathbb{R}^3} \frac{|\varphi(x)|^2}{|x-y|}dx \leq C\|\varphi\|_{L^2(\mathbb{R}^3)}\|\nabla\varphi\|_{L^2(\mathbb{R}^3)}$$

with a constant C independent of y and φ. Note that if $u, v \in H^1(\mathbb{R}^3)$ $\frac{u(x)v(y)}{\sqrt{|x-y|}} \in L^2(\mathbb{R}^3 \times \mathbb{R}^3)$. Indeed:

$$\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{u^2(x)v^2(y)}{|x-y|}dxdy = \int_{\mathbb{R}^3}\Big(\int_{\mathbb{R}^3}\frac{u^2(x)}{|x-y|}dx\Big)v^2(y)dy$$
$$\leq C\|u\|_{L^2(\mathbb{R}^3)}\|\nabla u\|_{L^2(\mathbb{R}^3)}\int_{\mathbb{R}^3} v^2(y)dy \leq C\|u\|_{L^2(\mathbb{R}^3)}\|\nabla u\|_{L^2(\mathbb{R}^3)}\|v\|^2_{L^2(\mathbb{R}^3)}$$

Let Ψ^m be a sequence weakly convergent in $\mathcal{H}$ to Ψ; this sequence is bounded in $\mathcal{H}$; without loss of generality it can be supposed that $\|\Psi^m\|_{\mathcal{H}} \leq 1$.
Consider a term of the form

$$\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{f(x)g(y)\Psi_i^m(x)\Psi_j^m(y)}{|x-y|}dxdy \tag{23}$$

where $f, g \in \{(\Phi_0)_1, ..., (\Phi_0)_N\}$. We have seen that $\frac{f(x)g(y)}{\sqrt{|x-y|}}$, $\frac{\Psi_i^m(x)\Psi_j^m(y)}{\sqrt{|x-y|}} \in$ $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$; since $\|\Psi^m\|_{\mathcal{H}} \leq 1$, it follows that $\frac{\Psi_i^m(x)\Psi_j^m(y)}{\sqrt{|x-y|}}$ is weakly

convergent in $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ to[2] $\frac{\Psi_i(x)\Psi_j(y)}{\sqrt{|x-y|}}$ so any term of the form (23) is weakly continuous (so also lower weakly semicontinuous), and of course the same is true for any sum of terms of this type, in particular $\frac{\rho_{\Phi_0,\Psi^m}(x)\rho_{\Phi_0,\Psi^m}(y)}{|x-y|}$, $\frac{\rho_{\Phi_0}(x,y)\rho_{\Psi^m}(x,y)}{|x-y|}$, $\frac{\rho_{\Phi_0,\Psi^m}(x,y)\rho_{\Phi_0,\Psi^m}(y,x)}{|x-y|}$.

The only term that remains to be analyzed in (23) is

$$4\iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{\rho_\Psi(x)\rho_{\Phi_0}(y) - \rho_{\Phi_0,\Psi}(x,y)^2}{|x-y|}dxdy$$

We transform the numerator of the above fraction as follows:

$$\begin{aligned}
&\rho_\Psi(x)\rho_{\Phi_0}(y) - (\rho_{\Phi_0,\Psi}(x,y))^2 = \sum_{i=1}^{N}(\Psi_i)^2(x)(\Phi_0)_i^2(y) \\
&+\sum_{i<j}(\Psi_i)^2(x)(\Phi_0)_j^2(y) + (\Psi_j)^2(x)(\Phi_0)_i^2(y) \\
&-\sum_{i=1}^{N}(\Psi_i)^2(x)(\Phi_0)_i^2(y) - \sum_{i<j}\Psi_i(x)(\Phi_0)_i(y)\Psi_j(x)(\Phi_0)_j(y) \\
&=\sum_{i<j}\Big(\Psi_i(x)(\Phi_0)_j(y) - \Psi_j(x)(\Phi_0)_i(y)\Big)^2
\end{aligned}$$

It is easy to see from this equality that $\rho_\Psi(x)\rho_{\Phi_0}(y) - (\rho_{\Phi_0,\Psi}(x,y))^2$ is a convex function of Ψ and therefore, by a classical functional analysis argument, is sequentially weakly lower semicontinuous. □

Proof of Proposition 1: Let us proceed with the proof of proposition 1. Suppose on the contrary that the conclusion is not true. Then there exists a sequence $\{\Psi^m\}_{m\geq 1} \in X_{\Phi_0}$ such that $\|\Psi^m\|_{\mathcal{H}} = 1$, and $\lim_{m\to\infty} a_{\Phi_0}(\Psi^m, \Psi^m) = 0$; extracting if necessary a subsequence out of it, we may suppose that $\{\Psi^m\}_{m\geq 1}$ is weakly convergent in $\mathcal{H}$ to

[2] In order to rigorously identify the weak limit one uses appropriate test functions $\sqrt{|x-y|}\alpha(x)\beta(y)\mathbf{1}_{|x|\leq R}\mathbf{1}_{|y|\leq R}$ for any $\alpha, \beta \in L^2(\mathbb{R}^3)$, $R > 0$.

$\Psi \in X_{\Phi_0}$. We first write:

$$a_{\Phi_0}(\Psi^m,\Psi^m) = 2\cdot\sum_{i=1}^{N}\int_{\mathbb{R}^3}|\nabla\Psi_i^m|^2 +$$
$$2\cdot\sum_{i,j=1}^{N}\Lambda_{ij}^0\int_{\mathbb{R}^3}\Psi_i^m\Psi_j^m + 2\cdot\sum_{i=1}^{N}\int_{\mathbb{R}^3}V(\Psi_i^m)^2 +$$
$$\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{8\rho_{\Phi_0,\Psi^m}(x)\rho_{\Phi_0,\Psi^m}(y)+4\rho_{\Psi^m}(x)\rho_{\Phi_0}(y)}{|x-y|}dxdy$$
$$-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{4\rho_{\Phi_0}(x,y)\rho_{\Psi^m}(x,y)}{|x-y|}dxdy$$
$$-\frac{1}{2}\iint_{\mathbb{R}^3\times\mathbb{R}^3}\frac{4\rho_{\Phi_0,\Psi^m}(x,y)(\rho_{\Phi_0,\Psi^m}(x,y)+\rho_{\Phi_0,\Psi^m}(y,x))}{|x-y|}dxdy \quad (24)$$

Recall that ([11] p.42) that $\int_{\mathbb{R}^3}V\psi^2dx$ is weakly lower semicontinuous on $H^1(\mathbb{R}^3)$ ([11] p.42). By the lemma 6 the integrals on $\mathbb{R}^3\times\mathbb{R}^3$ in (24) also have weakly lower semicontinuity properties. Since the matrix Λ^0 has **strictly positive** eigenvalues (remark 2) the first two terms on the right hand side of (24) define a norm so this part is also weakly lower semicontinuous ; we obtain

$$a_{\Phi_0}(\Psi,\Psi) \leq \lim_{m\to\infty} a_{\Phi_0}(\Psi^m,\Psi^m) = 0$$

which together with (1) imply $\Psi = 0$. We will use now this information for a finer analysis of the sequence $a_{\Phi_0}(\Psi^m,\Psi^m)$; by the argument above there exists a constant $c_0 > 0$ depending on Φ_0 such that for any $\Psi \in \mathcal{H}$:

$$\sum_{i=1}^{N}\int_{\mathbb{R}^3}2|\nabla\Psi_i|^2 + \sum_{i,j=1}^{N}\Lambda_{ij}^0\int_{\mathbb{R}^3}\Psi_i\Psi_j \geq c_0\|\Psi\|_{\mathcal{H}}.$$

Using again the lower semicontinuity of the remaining terms we obtain:

$$0 = \lim_{m\to\infty} a_{\Phi_0}(\Psi^m,\Psi^m) \geq 0 \;+\; \liminf_{m\to\infty}\sum_{i=1}^{N}\int_{\mathbb{R}^3}2|\nabla\Psi_i^m|^2 +$$
$$\sum_{i,j=1}^{N}\Lambda_{ij}^0\int_{\mathbb{R}^3}\Psi_i^m\Psi_j^m \geq c_0\liminf_{m\to\infty}\|\Psi^m\|_{\mathcal{H}} = c_0 > 0,$$

which is impossible. □

Motivated by the above analysis, we will introduce the following hypothesis:

$$\forall \Psi \in \Phi_0^{\perp\!\!\!\perp}, \Psi \neq 0 : \; a_{\Phi_0}(\Psi, \Psi) > 0. \tag{25}$$

which, by proposition 1, ensures the existence of a "coercivity constant" $c_{\Phi_0} > 0$ such that

$$\forall \Psi \in \Phi_0^{\perp\!\!\!\perp}, \Psi \neq 0 : \; a_{\Phi_0}(\Psi, \Psi) \geq c_{\Phi_0} \|\Psi\|_{\mathcal{H}}^2. \tag{26}$$

Remark 6 Using the lemma 5 a posteriori analysis may still be carried out without the hypothesis 25 ; some aspects of a more general analysis are presented in remark 11.

4 Error estimators, bounds and convergence acceleration

Let $\Phi_0, \Phi \in \mathcal{H} \cap \mathcal{K}$ be as in section 2.1: Φ_0 a minimizer of (8) (which is thus a solution of (4)) and $\Phi \in \mathcal{H} \cap \mathcal{K}$ a given discrete approximation of Φ_0 obtained by a previous computation.

Let us denote by $\epsilon = \|U_{\Phi_0,\Phi}\Phi_0 - \Phi\|_{\mathcal{H}} = \|U_{\Phi,\Phi_0}\Phi - \Phi_0\|_{\mathcal{H}}$ the distance between Φ and Φ_0. Even if the wavefunction Φ_0 may be intrinsically interesting (e.g. when the form of the molecular orbitals is studied), the main result of a Hartree-Fock computation is the Hartree-Fock energy $\mathcal{E}^{HF}(\Phi_0)$.

We will suppose in all that follows that Φ is close enough to Φ_0 such that e.g. in the development of the error $\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0)$ with respect to powers of ϵ: $\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0) = c_k\epsilon^k + o(\epsilon^k)$ the second term $o(\epsilon^k)$ is indeed smaller than $c_k\epsilon^k$ (due to the asymptotic properties of the decomposition this is certain to happen when ϵ is small enough).

4.1 Error estimators

The a posteriori analysis method presented in this section is connected to the works of Babuška [1], Bernardi [4], Ladevèze [9], Oden [14], Pousin and Rappaz [18], Verfürth [21,22] and is aimed at giving quantitative indications on the form of the error, through **bilateral** estimates. Even if the constants are not explicitly known, this method may prove interesting when only relative error estimates are needed (as in adaptative procedures) or when the estimator is shown to possess further properties that allow to estimate those constants.

Let us recall (see also (15)) that $U_{\Phi,\Phi_0}\Phi - \Phi_0 \in \mathcal{S}_{\Phi_0} \oplus \Phi_0^{\perp\!\!\!\perp}$ and denote $U_{\Phi,\Phi_0}\Phi - \Phi_0 = S\Phi_0 + W$, $S\Phi_0 \in \mathcal{S}_{\Phi_0}$, $W \in \Phi_0^{\perp\!\!\!\perp}$. Then one can write

$$\begin{aligned}
&\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0) = \mathcal{E}^{HF}(U_{\Phi,\Phi_0}\Phi) - \mathcal{E}^{HF}(\Phi_0) \\
&= \mathcal{E}^{\Phi_0}(\Phi_0 + S\Phi_0 + W) - \mathcal{E}^{\Phi_0}(\Phi_0) \\
&= D\mathcal{E}^{\Phi_0}(\Phi_0)(S\Phi_0 + W) + D^2\mathcal{E}^{\Phi_0}(\Phi_0)(S\Phi_0 + W, S\Phi_0 + W) + O(\epsilon^3) \\
&= 0 + D^2\mathcal{E}^{\Phi_0}(\Phi_0)(W,W) + O(\epsilon^3) = a_{\Phi_0}(W,W) + O(\epsilon^3)
\end{aligned}$$

where we have used firstly the fact that Φ_0 is the solution of (4) (see remark 5 equation 21) and secondly the lemma 2 for $(U_{\Phi,\Phi_0}\Phi, \Phi_0) \to (\Psi, \Phi)$. From the continuity of a_{Φ_0} and (26) one concludes that $\|W\|^2_{\mathcal{H}}$ is a third order estimator of the energy error $\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0)$.

Remark 7 It easy to see by (19) that $\|W\|_{\mathcal{H}} = \epsilon + O(\epsilon^2)$.

Unfortunately direct computation of W (and then of $\|W\|^2_{\mathcal{H}}$) assumes knowledge of Φ_0 which is not available. However good approximations of $\|W\|^2_{\mathcal{H}}$ that require only the knowledge of Φ can be found. Indeed, let us set $F = D\mathcal{E}^{HF}$, $\Psi = U_{\Phi,\Phi_0}\Phi$ and study the norm of $F(\Psi)$ in the dual space $\Psi^{\perp\!\!\!\perp *}$ of $\Psi^{\perp\!\!\!\perp}$

$$\begin{aligned}
&\|F(\Psi)\|_{\Psi^{\perp\!\!\!\perp *}} = \sup_{\xi \in \Psi^{\perp\!\!\!\perp}} \frac{< D\mathcal{E}^{HF}(\Psi), \xi >}{\|\xi\|_{\mathcal{H}}} = \sup_{\xi \in \Psi^{\perp\!\!\!\perp}} \frac{< D\mathcal{E}^{\Phi_0}(\Psi), \xi >}{\|\xi\|_{\mathcal{H}}} \\
&= \sup_{\xi \in \Psi^{\perp\!\!\!\perp}} \frac{< D\mathcal{E}^{\Phi_0}(\Psi) - D\mathcal{E}^{\Phi_0}(\Phi_0), \xi >}{\|\xi\|_{\mathcal{H}}} \\
&= \sup_{\xi \in \Psi^{\perp\!\!\!\perp}} \frac{D^2\mathcal{E}^{\Phi_0}(\Phi_0)(\Psi - \Phi_0, \xi)}{\|\xi\|_{\mathcal{H}}} + O(\epsilon^2)
\end{aligned}$$

We used in the first line of the equation above the definition (20) of $\mathcal{E}^{\Phi_0}$ and the identity $DF_{ij}(\Psi, \cdot) \equiv 0$ on $\Psi^{\perp\!\!\!\perp}$. We show now that we can replace in the above supremum the space $\Psi^{\perp\!\!\!\perp} = (U_{\Phi,\Phi_0}\Phi)^{\perp\!\!\!\perp} = \Phi^{\perp\!\!\!\perp}$ by $\Phi_0^{\perp\!\!\!\perp}$. Let $\xi \in \Psi^{\perp\!\!\!\perp}$ be written as $\xi = M\Phi_0 + \tilde{\xi}$, $\tilde{\xi} \in \Phi_0^{\perp\!\!\!\perp}$. Note that

$$\begin{aligned}
&|M_{ij}| = | < \xi_i, \Phi_{0j} > | = | < \xi_i, \Phi_{0j} - \Psi_j > | \\
&\leq \|\xi\|_{(L^2(\mathbb{R}^3))^N} \|\Phi_0 - \Psi\|_{(L^2(\mathbb{R}^3))^N}
\end{aligned} \tag{27}$$

so one can write

$$|\frac{a_{\Phi_0}(\Psi - \Phi_0, M\Phi_0)}{\|\xi\|_{\mathcal{H}}}| \leq \frac{C_{\Phi_0}\|\Psi - \Phi_0\|_{\mathcal{H}}\|\xi\|_{\mathcal{H}}\|\Phi_0 - \Psi\|_{(L^2(\mathbb{R}^3))^N}}{\|\xi\|_{\mathcal{H}}} \leq C_{\Phi_0}\epsilon^2,$$

where C_{Φ_0} is the continuity constant of a_{Φ_0}. Since $\frac{\|\tilde{\xi}\|_{\mathcal{H}}}{\|\xi\|_{\mathcal{H}}} = 1 + O(\epsilon)$ one concludes that

$$\|F(\Psi)\|_{\Psi^{\perp\!\!\!\perp *}} = \sup_{\xi \in \Psi^{\perp\!\!\!\perp}} \frac{a_{\Phi_0}(\Psi - \Phi_0, \tilde{\xi})}{\|\tilde{\xi}\|_{\mathcal{H}}} + O(\epsilon^2)$$
$$= \sup_{\tilde{\xi} \in \Psi_0^{\perp\!\!\!\perp}} \frac{a_{\Phi_0}(\Psi - \Phi_0, \tilde{\xi})}{\|\tilde{\xi}\|_{\mathcal{H}}} + O(\epsilon^2) = \sup_{\tilde{\xi} \in \Psi_0^{\perp\!\!\!\perp}} \frac{a_{\Phi_0}(S\Phi_0 + W, \tilde{\xi})}{\|\tilde{\xi}\|_{\mathcal{H}}} + O(\epsilon^2)$$
$$= \sup_{\tilde{\xi} \in \Psi_0^{\perp\!\!\!\perp}} \frac{a_{\Phi_0}(W, \tilde{\xi})}{\|\tilde{\xi}\|_{\mathcal{H}}} + O(\epsilon^2) = \|W\|_{\mathcal{H}} + O(\epsilon^2).$$

We have shown above that $\|F(\Psi)\|_{\Psi^{\perp\!\!\!\perp *}}$ is a second order approximation of $\|W\|_{\mathcal{H}}$ and therefore $\|F(\Psi)\|^2_{\Psi^{\perp\!\!\!\perp *}}$ will be a third order estimator of the energy error $\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0)$. We next prove that $\|F(\Psi)\|_{\Psi^{\perp\!\!\!\perp *}}$ is invariant with respect to the multiplication of Ψ by unitary matrices and therefore equal to $\|F(\Phi)\|_{\Phi^{\perp\!\!\!\perp *}}$, so it can be computed (a posteriori) using only available data (i.e. Φ). Let us compute for ζ in $\mathcal{H} \cap \mathcal{K}$ the function $F(U\zeta)$, by the definition of F this equals $D\mathcal{E}^{HF}(U\zeta)$ which can be written:

$$D\mathcal{E}^{HF}(U\zeta) = \Big(\mathcal{F}_{U\zeta}((U\zeta)_i)\Big)_{i=1}^N = \Big((-\frac{1}{2}\Delta + V)((U\zeta)_i)\Big)_{i=1}^N +$$
$$\Big((\rho_{U\zeta} \star \frac{1}{|x|})(U\zeta)_i - \int_{\mathbb{R}^3} \frac{\rho_{U\zeta}(x,y)}{|x-y|}(U\zeta)_i(y)dy\Big)_{i=1}^N$$
$$= U\Big((-\frac{1}{2}\Delta + V)(\zeta_i)\Big)_{i=1}^N + U\Big((\rho_\zeta \star \frac{1}{|x|})\zeta_i - \int_{\mathbb{R}^3} \frac{\rho_\zeta(x,y)}{|x-y|}\zeta_i(y)dy\Big)_{i=1}^N,$$

where we have used the invariance property (6). It was therefore proven that

$$F(U\zeta) = UF(\zeta), \forall \zeta \in \mathcal{H} \cap \mathcal{K},$$

and therefore $\|F(\Psi)\|_{\Psi^{\perp\!\!\!\perp *}} = \|F(U_{\Phi,\Phi_0}\Phi)\|_{\Psi^{\perp\!\!\!\perp *}} = \|F(\Phi)\|_{\Phi^{\perp\!\!\!\perp *}}$ We will summarize the results obtained in this section in the following

Theorem 1 *Let Φ_0 be a minimizer of (8), $\Phi \in \mathcal{H} \cap \mathcal{K}$ a (given) discrete approximation of Φ_0 obtained by a previous computation as described in section 2.1 (10), and denote $\epsilon = \|U_{\Phi_0,\Phi}\Phi_0 - \Phi\|_{\mathcal{H}}$ the quotient distance between Φ and Φ_0. Then, under the assumption (25),*

$$\|D\mathcal{E}^{HF}(\Phi)\|_{\Phi^{\perp\!\!\!\perp *}} = \epsilon + O(\epsilon^2).$$

Moreover there exists constants c_1, c_2 depending only on Φ_0 such that

$$c_1\|D\mathcal{E}^{HF}(\Phi)\|^2_{\Phi^{\perp\!\!\!\perp *}} + O(\epsilon^3) \leq \mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0)$$
$$\leq c_2\|D\mathcal{E}^{HF}(\Phi)\|^2_{\Phi^{\perp\!\!\!\perp *}} + O(\epsilon^3). \qquad (28)$$

Remark 8 The constants c_1, c_2 in (28) are not known and therefore the quantity $\|D\mathcal{E}^{HF}(\Phi)\|^2_{\Phi^{\perp\!\!\!\perp}*}$ can be used to **estimate** the error in energy but not to obtain precise error bars.

4.2 Explicit bounds for the Hartree-Fock energy and convergence acceleration

The purpose of this section is to propose methods to find explicit bounds for the Hartree-Fock energy. The method belongs to the more general paradigm [13,15–17] of definition of explicit lower and upper bounds for outputs depending on the solution of a partial differential equation. The output of interest will be taken to be the Hartree-Fock energy ; this choice will be seen (cf. thm. 2) to posses particularities that in fact allow to design an improvement of the solution itself, although this is not expected to be the case for general outputs.

We will begin this section with some remarks on the coercivity properties of the bilinear forms a_{Φ_0} and a_Φ.

Lemma 7 *Under the hypothesis (25) there exists a constant $\gamma > 0$ depending only on Φ_0 such that for any $U \in \mathcal{U}(N)$ the bilinear form $a_{U\Phi_0}$ is coercive on $(U\Phi_0)^{\perp\!\!\!\perp} = \Phi_0^{\perp\!\!\!\perp}$ with coercivity constant γ.*

Proof. Note that for any $\Psi_1 \in \mathcal{H} \cap \mathcal{K}$, $\Psi_2 \in \mathcal{H}$, $U \in \mathcal{U}(N)$: $a_{U\Psi_1}(U\Psi_2, U\Psi_2) = a_{\Psi_1}(\Psi_2, \Psi_2)$, so by (25) and proposition 1 we obtain the conclusion. □

Lemma 8 *Under the assumption (25) there exists a constant $\eta > 0$ depending only on Φ_0 such that for all $\Phi \in \mathcal{H} \cap \mathcal{K}$ with $\|\Phi - \Phi_0\|_{\mathcal{H}} \leq \eta$ the bilinear form a_Φ is coercive on $\Phi^{\perp\!\!\!\perp}$ with a coercivity constant depending only of Φ_0.*

Proof. Let $\xi \in \Phi^{\perp\!\!\!\perp}$, $\|\xi\|_{\mathcal{H}} \leq 1$ be written as $\xi = M\Phi_0 + \tilde{\xi}$, $\tilde{\xi} \in \Phi_0^{\perp\!\!\!\perp}$. We will generically denote by C various constants depending only on Φ_0. Recall that (by (27)) $|M_{ij}| \leq \|\xi\|_{(L^2(\mathbb{R}^3))^N} \|\Phi_0 - \Psi\|_{(L^2(\mathbb{R}^3))^N}$, so for $\|\Phi_0 - \Psi\|_{\mathcal{H}}$ small enough

$$a_\Phi(\xi, \xi) = a_\Phi(\tilde{\xi} + M\Phi_0, \tilde{\xi} + M\Phi_0) \geq a_\Phi(\tilde{\xi}, \tilde{\xi}) - C\|\xi\|_{\mathcal{H}}\|\tilde{\xi}\|_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}}.$$

But for $\|\Phi_0 - \Psi\|_{\mathcal{H}}$ small enough we can also write

$$\begin{aligned}\|\xi\|_{\mathcal{H}}\|\tilde{\xi}\|_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}} &\leq \|\xi\|_{\mathcal{H}}(\|\xi\|_{\mathcal{H}} + \|\xi\|_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}})\|\Phi_0 - \Psi\|_{\mathcal{H}} \\ &\leq C\|\xi\|^2_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}}.\end{aligned}$$

Since $|\Lambda_{ij} - \Lambda^0_{ij}| \leq C\|\Phi - \Phi_0\|_{\mathcal{H}}$ it follows that $|a_\Phi(\tilde{\xi},\tilde{\xi}) - a_{\Phi_0}(\tilde{\xi},\tilde{\xi})| \leq C\|\tilde{\xi}\|^2_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}}$ so in fact

$$\begin{aligned} a_\Phi(\xi,\xi) &\geq a_{\Phi_0}(\tilde{\xi},\tilde{\xi}) - C(\|\tilde{\xi}\|^2_{\mathcal{H}} + \|\xi\|^2_{\mathcal{H}})\|\Phi_0 - \Psi\|_{\mathcal{H}} \\ &\geq \gamma\|\tilde{\xi}\|^2_{\mathcal{H}} - C(\|\tilde{\xi}\|^2_{\mathcal{H}} + \|\xi\|^2_{\mathcal{H}})\|\Phi_0 - \Psi\|_{\mathcal{H}}. \end{aligned}$$

It suffices now to use a last time $|\|\xi\|_{\mathcal{H}} - \|\tilde{\xi}\|_{\mathcal{H}}| \leq \|\xi\|_{\mathcal{H}}\|\Phi_0 - \Psi\|_{\mathcal{H}}$ to conclude. □

In what follows we start the presentation of the construction of (lower) bounds for the Hartree-Fock energy. As it was seen in lemma 7, under the assumption (25) we have uniform coercivity properties for bilinear forms a_{Φ_0} with respect to the multiplication of Φ_0 by unitary matrices $U \in \mathcal{U}(N)$; for this reason we can replace Φ_0 with any $U\Phi_0$ that fits better our needs; we will therefore suppose in agreement with lemma 1 that Φ_0 is such that $\Phi_0 - \Phi = S\Phi + W \in \mathcal{S}_\Phi \oplus \Phi^{\perp\!\!\!\perp}$.

The construction of (lower) bounds for the Hartree-Fock energy is based on the following development:

$$\begin{aligned} \mathcal{E}^{HF}(\Phi_0) - \mathcal{E}^{HF}(\Phi) &= \mathcal{E}^\Phi(\Phi_0) - \mathcal{E}^\Phi(\Phi) = \mathcal{E}^\Phi(\Phi + S\Phi + W) - \mathcal{E}^\Phi(\Phi) \\ &= D\mathcal{E}^\Phi(\Phi)(S\Phi + W) + \frac{1}{2}D^2\mathcal{E}^\Phi(\Phi)(S\Phi + W, S\Phi + W) + O(\epsilon^3) \end{aligned}$$

Note first that by the properties of Φ as described in section 2.1 eq. (10) $D\mathcal{E}^\Phi(\Phi)$ is null on the dual space of the discretization space so in particular $D\mathcal{E}^\Phi(\Phi)(S\Phi) = 0$; recall also the fact that $S\Phi$ is of order ϵ^2 and W of order ϵ to obtain

$$\begin{aligned} \mathcal{E}^{HF}(\Phi_0) - \mathcal{E}^{HF}(\Phi) = D\mathcal{E}^\Phi(\Phi)(W) + \\ \frac{1}{2}D^2\mathcal{E}^\Phi(\Phi)(W, W) + O(\epsilon^3) \end{aligned} \tag{29}$$

Consider now the problem: find the *reconstructed error* $\hat{W} \in \Phi^{\perp\!\!\!\perp}$ such that

$$D^2\mathcal{E}^\Phi(\Phi)(\hat{W}, \Psi) + D\mathcal{E}^\Phi(\Phi)(\Psi) = 0, \ \forall \Psi \in \Phi^{\perp\!\!\!\perp}. \tag{30}$$

By the coercivity of a_Φ it follows that (30) has a unique solution $\hat{W} \in \Phi^{\perp\!\!\!\perp}$.

Remark 9 Note that in order to compute $\hat{W}$ one solves a **direct** (i.e. not eigenvalue) problem on the solution space ; moreover all operators involved depend only on Φ.

Using the definition of $\hat{W}$ one can rewrite (29):

$$\mathcal{E}^{HF}(\Phi_0) = \mathcal{E}^{HF}(\Phi) - D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, W) + \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(W, W)$$
$$+O(\epsilon^3) = \mathcal{E}^{HF}(\Phi) - \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W})$$
$$+\frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(W - \hat{W}, W - \hat{W}) + O(\epsilon^3). \quad (31)$$

But since a_{Φ} is positive on $\Phi^{\perp\!\!\!\perp}$ it follows that $\frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(W-\hat{W}, W-\hat{W}) \geq 0$ so in fact we obtain an asymptotic **explicit lower bound on the Hartree-Fock energy**:

$$\mathcal{E}^{HF}(\Phi_0) \geq \mathcal{E}^{HF}(\Phi) - \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}) + O(\epsilon^3), \quad (32)$$

which together with the inequality $\mathcal{E}^{HF}(\Phi_0) \leq \mathcal{E}^{HF}(\Phi)$ gives an interval for the **exact** value of the Hartree-Fock energy.

Remark 10 A natural question is to study the order in ϵ of the length of the error bar found above. Let us recall that the error in energy is of order ϵ^2 ; we will prove that this interval is optimal in a sense that its length is also of order ϵ^2 ; indeed the distance between the upper and lower bound is $\frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}) + O(\epsilon^3)$ which is equivalent to $\|\hat{W}\|_{\mathcal{H}}$; all that remains to be proven is that $\|\hat{W}\|_{\mathcal{H}} \leq C\epsilon$ (with a constant not depending on Φ_0). Indeed:

$$\|\hat{W}\|_{\mathcal{H}} \leq C\|D\mathcal{E}^{\Phi}(\Phi)\|_{\Phi^{\perp\!\!\!\perp *}} \leq C\|D\mathcal{E}^{\Phi}(\Phi) - D\mathcal{E}^{\Phi}(\Phi_0)\|_{\Phi^{\perp\!\!\!\perp *}}$$
$$+C\|D\mathcal{E}^{\Phi}(\Phi_0) - D\mathcal{E}^{\Phi_0}(\Phi_0)\|_{\Phi^{\perp\!\!\!\perp *}} \leq C\epsilon$$

where we have used the fact that $D\mathcal{E}^{\Phi_0}(\Phi_0)$ is null on $\Phi_0^{\perp\!\!\!\perp}$.

The nomination of $\hat{W}$ as "reconstructed error" is best explained by the following property:

$$\hat{W} = W + O(\epsilon^2). \quad (33)$$

In order to prove (33) we will prove that W has the following property:

$$|D^2\mathcal{E}^{\Phi}(\Phi)(W, \Psi) + D\mathcal{E}^{\Phi}(\Phi)(\Psi)| \leq C\epsilon^2, \ \forall\Psi \in \Phi^{\perp\!\!\!\perp}, \ \|\Psi\|_{\mathcal{H}} = 1. \quad (34)$$

with a constant C independent of Φ, Ψ. Suppose (34) is true then jointly with (30) one obtains:

$$|D^2\mathcal{E}^{\Phi}(\Phi)(W - \hat{W}, \Psi)| \leq C\epsilon^2, \forall\Psi \in \Phi^{\perp\!\!\!\perp}, \ \|\Psi\|_{\mathcal{H}} = 1.$$

Let $\Psi = \frac{W-\hat{W}}{\|W-\hat{W}\|_{\mathcal{H}}}$; from the coercivity of $a_\Phi = D^2\mathcal{E}^\Phi(\Phi)$ we deduce:

$$\frac{1}{\|W-\hat{W}\|_{\mathcal{H}}} \cdot c\|W-\hat{W}\|_{\mathcal{H}}^2 \leq C\epsilon^2,$$

and (33) follows.

Recall that, from lemma 2, $\|\Phi_0 - \Phi - W\|$ is of order ϵ^2. In order to prove (34) it is thus sufficient to prove it for $\Phi_0 - \Phi$ instead of W: let us write

$$D\mathcal{E}^\Phi(\Phi)(\Psi) = D\mathcal{E}^\Phi(\Phi_0)(\Psi) + D^2\mathcal{E}^\Phi(\Phi_0)(\Phi-\Phi_0,\Psi) + O(\epsilon^2).$$

Besides we have

$$|D^2\mathcal{E}^\Phi(\Phi_0)(\Phi-\Phi_0,\Psi) - D^2\mathcal{E}^\Phi(\Phi)(\Phi-\Phi_0,\Psi)| \leq C\epsilon^2\|\Psi\|_{\mathcal{H}},$$

(with a constant C depending only of Φ_0), so

$$D\mathcal{E}^\Phi(\Phi)(\Psi) = D\mathcal{E}^\Phi(\Phi_0)(\Psi) + D^2\mathcal{E}^\Phi(\Phi)(\Phi-\Phi_0,\Psi) + O(\epsilon^2)$$

and therefore

$$D^2\mathcal{E}^\Phi(\Phi)(\Phi_0-\Phi,\Psi) + D\mathcal{E}^\Phi(\Phi)(\Psi) = D\mathcal{E}^\Phi(\Phi_0)(\Psi) + O(\epsilon^2).$$

It suffices now to prove that $D\mathcal{E}^\Phi(\Phi_0)(\Psi) = O(\epsilon^2)$. By the definition of $\mathcal{E}^\Phi$,

$$\begin{aligned} D\mathcal{E}^\Phi(\Phi_0)(\Psi) &= D\mathcal{E}^{\Phi_0}(\Phi_0)(\Psi) + \textstyle\sum_{i,j=1}^N(\Lambda_{ij}-\Lambda_{ij}^0)DF_{ij}(\Phi_0)(\Psi) \\ &= 0 + \textstyle\sum_{i,j=1}^N(\Lambda_{ij}-\Lambda_{ij}^0)DF_{ij}(\Phi_0)(\Psi). \end{aligned}$$

Note firstly that $\Lambda_{ij} - \Lambda_{ij}^0 \leq C\epsilon$ (C depending only of Φ_0). Moreover

$$\begin{aligned} DF_{ij}(\Phi_0)(\Psi) &= <\Phi_{0i},\Psi_j> + <\Phi_{0j},\Psi_i> \\ &= <\Phi_{0i}-\Phi_i,\Psi_j> + <\Phi_{0j}-\Phi_j,\Psi_i> \end{aligned}$$

thus $|DF_{ij}(\Phi_0)(\Psi)|$ can be upper bounded by $C\epsilon$ (we used the fact that $\Psi \in \Phi^{\perp\!\!\!\perp}$), which concludes the proof of (33).

Combining (31) and (33) we can give a better version of (32):

$$\mathcal{E}^{HF}(\Phi_0) = \mathcal{E}^{HF}(\Phi) - \frac{1}{2}D^2\mathcal{E}^\Phi(\Phi)(\hat{W},\hat{W}) + O(\epsilon^3), \tag{35}$$

so instead of a lower bound we have obtained an **improvement** of the Hartree-Fock energy ; note that this improvement is of a **strictly** higher order in ϵ since the best approximation known before the computation of $\hat{W}$ was $\mathcal{E}^{HF}(\Phi)$ which is exact to the order ϵ^2.

Although (35) may represent in itself the conclusion of the a posteriori analysis, further progress is possible. To this end note that

an improvement for the wavefunction Φ has also been found, namely $\tilde{\Phi} = \Phi + \hat{W}$. However we cannot propose $\tilde{\Phi}$ as a legitimate solution of (4) since it is not certain to be in $\mathcal{K}$. We will see in the following that it is possible to find a correction to add to $\Phi + \hat{W}$ which not only gives an admissible solution of (4) but also **improves with another order** the approximation (35) of the Hartree-Fock energy $\mathcal{E}^{HF}(\Phi_0)$.

In order to improve even more the solution, remind the equality $\Phi_0 = \Phi + W + S\Phi$. Since both Φ_0 and Φ are in $\mathcal{K}$ we can write

$$\delta_{ij} =< \Phi_{0i}, \Phi_{0j} >=< \Phi_i + \sum_{k=1}^{N} S_{ik}\Phi_k + W_i, \Phi_j + \sum_{l=1}^{N} S_{jl}\Phi_l + W_j >$$

$$= \delta_{ij} + < W_i, W_j > + \sum_{k=1}^{N} S_{ik}\delta_{kj} + \sum_{k=1}^{N} S_{jl}\delta_{il} + O(\epsilon^4) \tag{36}$$

because we know that $S_{ij} = O(\epsilon^2)$. We obtain

$$0 =< W_i, W_j > + S_{ij} + S_{ji} + O(\epsilon^4) =< \hat{W}_i, \hat{W}_j > + S_{ij} + S_{ji} + O(\epsilon^3)$$

so denoting $\tilde{S}_{ij} = -\frac{1}{2} < \hat{W}_i, \hat{W}_j >$, we obtain that $\tilde{S}\Phi$ is a order ϵ^3 approximation of $S\Phi$: $\tilde{S}\Phi = S\Phi + O(\epsilon^3)$. Note that by remark 9 that the computation of $\tilde{S}$ requires knowledge of Φ only.

We will prove in the following that having an approximation $\hat{W}$ of W to the order ϵ^2 and an approximation $\tilde{S}$ of S to the order ϵ^3 is enough to have an approximation of the Hartree-Fock energy to the order ϵ^4. Indeed, write

$$\mathcal{E}^{HF}(\Phi_0) - \mathcal{E}^{HF}(\Phi) = \mathcal{E}^{\Phi}(\Phi_0) - \mathcal{E}^{\Phi}(\Phi) = \mathcal{E}^{\Phi}(\Phi + S\Phi + W) - \mathcal{E}^{\Phi}(\Phi)$$

$$= D\mathcal{E}^{\Phi}(\Phi)(S\Phi + W) + \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(S\Phi + W, S\Phi + W)$$

$$+\frac{1}{3!}D^3\mathcal{E}^{\Phi}(\Phi)(S\Phi + W, S\Phi + W, S\Phi + W) + O(\epsilon^4)$$

$$= D\mathcal{E}^{\Phi}(\Phi)(W) + \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(W, W) + D^2\mathcal{E}^{\Phi}(\Phi)(S\Phi, W)$$

$$+\frac{1}{3!}D^3\mathcal{E}^{\Phi}(\Phi)(W, W, W) + O(\epsilon^4)$$

$$= -\frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}) + \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(W - \hat{W}, W - \hat{W}) +$$

$$D^2\mathcal{E}^{\Phi}(\Phi)(\tilde{S}\Phi, \hat{W}) + \frac{1}{3!}D^3\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}, \hat{W}) + O(\epsilon^4)$$

$$= -\frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}) + D^2\mathcal{E}^{\Phi}(\Phi)(\tilde{S}\Phi, \hat{W}) +$$

$$\frac{1}{3!}D^3\mathcal{E}^{\Phi}(\Phi)(\hat{W}, \hat{W}, \hat{W}) + O(\epsilon^4),$$

so we have obtained

$$\mathcal{E}^{HF}(\Phi_0) = \mathcal{E}^{HF}(\Phi) - \frac{1}{2}D^2\mathcal{E}^{\Phi}(\Phi)(\hat{W},\hat{W}) + D^2\mathcal{E}^{\Phi}(\Phi)(\tilde{S}\Phi,\hat{W}) + \frac{1}{3!}D^3\mathcal{E}^{\Phi}(\Phi)(\hat{W},\hat{W},\hat{W}) + O(\epsilon^4).$$

where all terms involved in the right hand side can be computed from Φ.

One problem remains though, our best approximation for the solution Φ_0, namely $\tilde{\tilde{\Phi}} = \Phi + \hat{W} + \tilde{S}\Phi$ is still not certain to be in $\mathcal{K}$; in fact it can be proved that there exists an $\hat{S}$ that depends only of Φ that has the property $\hat{S}\Phi = \tilde{S}\Phi + O(\epsilon^3)$ and such that $\hat{\Phi} = \Phi + \hat{W} + \hat{S}\Phi \in \mathcal{K}$. Moreover, using the above arguments, we will also have $\mathcal{E}^{HF}(\Phi_0) = \mathcal{E}^{HF}(\hat{\Phi}) + O(\epsilon^4)$. The existence and properties of $\hat{S}$ follows by considering as in (36) the equations satisfied by $\hat{S}$. Denote by M the matrix with entries $< \hat{W}_i, \hat{W}_j >$ then $\hat{S}$ is solution of the equation

$$(I + \hat{S})^2 = I - M. \tag{37}$$

This shows that $\hat{S}$ is an $O(\epsilon^3)$ approximation of $\tilde{S}$. The matrix $\hat{S}$ can be computed from equation 37 by taking the square root of $I - M$ which is well defined as $\hat{W}$ is close to W (and small). Note that this procedure may be costly for non-sparse matrices and can be replaced in practice with Taylor-like series expansion formulas

$$I + \hat{S} = \sqrt{I - M} = I - \frac{1}{2}M + \frac{1}{8}M^2 - \frac{1}{16}M^3 + ...$$

We will summarize the results obtained in this section in the following theorem:

Theorem 2 *Let Φ_0 be a minimizer of (8), $\Phi \in \mathcal{H} \cap \mathcal{K}$ a (known) discrete approximation of Φ_0 obtained by a previous computation as described in section 2.1 (10). Then, under the assumption (25), there exists an $\eta > 0$ such that for any $\Phi \in \mathcal{H} \cap \mathcal{K}$ with $\|U_{\Phi_0,\Phi}\Phi_0 - \Phi\| \leq \eta$ there exists $\hat{W} \in \Phi^{\perp\!\!\!\perp}$ and $\hat{S}\Phi \in \mathcal{S}_\Phi$ whose computation requires only knowledge of Φ such that $\hat{\Phi} = \Phi + \hat{S}\Phi + \hat{W} \in \mathcal{H} \cap \mathcal{K}$ has the following properties:*

$$\|\hat{\Phi} - \Phi_0\|_{\mathcal{H}} \leq c_1\|\Phi - \Phi_0\|^2_{\mathcal{H}},$$
$$|\mathcal{E}^{HF}(\hat{\Phi}) - \mathcal{E}^{HF}(\Phi_0)| \leq c_2|\mathcal{E}^{HF}(\Phi) - \mathcal{E}^{HF}(\Phi_0)|^2.$$

with constants c_1,c_2 depending only of Φ_0.

Error bars can be easily derived from the Thm. 2 above and the minimization properties of $\mathcal{E}^{HF}(\Phi_0)$:

Theorem 3 *Under the same assumptions and with the same notations as in Thm. 2, there exists an $\tilde{\eta} > 0$ such that for any $\Phi \in \mathcal{H} \cap \mathcal{K}$ with $\|U_{\Phi_0,\Phi}\Phi_0 - \Phi\| \leq \tilde{\eta}$ the following estimates hold:*

$$2 \cdot \mathcal{E}^{HF}(\hat{\Phi}) - \mathcal{E}^{HF}(\Phi) \leq \mathcal{E}^{HF}(\Phi_0) \leq \mathcal{E}^{HF}(\Phi).$$

Remark 11 The approach described in this section can be developed under more general assumptions than (25). Denote by X_{Φ_0} the closed subspace of $\Phi_0^{\perp\!\!\!\perp}$ where (1) holds so that, in agreement with proposition 1 a_{Φ_0} is coercive on X_{Φ_0} ; using the same arguments as in lemma 8 one proves for $\|\Phi_0 - \Phi\|_{\mathcal{H}}$ small enough coercivity for a_Φ on $X_{\Phi_0} \cap \Phi^{\perp\!\!\!\perp}$; this shows that the problem (30) has an unique solution on $X_{\Phi_0} \cap \Phi^{\perp\!\!\!\perp}$ and this solution is then shown to posses the same property (33) as $\hat{W}$. A "reconstructed symmetrical" part is then computed by the same method as above and we obtain thus an improvement for the energy and for the wavefunction. The only computational impediment to this program is that one cannot really identify the space $X_{\Phi_0} \cap \Phi^{\perp\!\!\!\perp}$ where problem (30) is to be solved ; one chooses then the largest subspace in $\Phi^{\perp\!\!\!\perp}$ where a_Φ is positive (therefore coercive), which will contain $X_{\Phi_0} \cap \Phi^{\perp\!\!\!\perp}$, and proves that the solution of (30) on this space is an order ϵ^2 approximation of the solution of (30) on $X_{\Phi_0} \cap \Phi^{\perp\!\!\!\perp}$. In practice (cf. section 5) there was no need to implement this procedure as (25) seems to be satisfied.

Remark 12 The numerical computation of $\hat{W}$ involves the resolution of equation (30) over the discrete subspace $\Phi_\delta^{\perp\!\!\!\perp}$ of $\Phi^{\perp\!\!\!\perp}$; the corresponding solution $\hat{W}_\delta$ will be an approximation of $\hat{W}$ which converges to $\hat{W}$ when the discretization parameter δ is such that $\Phi_\delta^{\perp\!\!\!\perp}$ converges to the space $\Phi^{\perp\!\!\!\perp}$.

Remark 13 Upon writing this paper we were made aware [5] that (30) is equivalent to a density matrix quadratic convergence equation (see for instance [3] an references therein for an introduction). A study is being undertaken to further investigate the advantages that this equivalence may bring at the numerical level.

5 Numerical simulations

The theory presented in the previous sections was tested in two categories of numerical experiments.

In the experiments of the first category we checked on simple cases (hydrogen molecule, helium) that the methodology proposed above is coherent with available results when the problem (30) that provides $\hat{W}$ is solved on a very fine discretization of $\mathcal{H}$.

In a second stage more complex molecules were studied and the method was implemented in a Hartree-Fock quantum chemistry code.

Before presenting the results let us remark that the partial differential equation (PDE) (30) is, for N large, very difficult to discretize with classical tools from the PDE equations (finite elements, finite volumes, ...) due to the high dimensionality of the linear spaces involved. Moreover a good discretization has also to take into account some specific quantum chemical effects as the singularities of the electronic wavefunction around nuclei; in conclusion, only very small quantum systems are thus available for study using classical tools in solving PDEs ; such systems are for example the hydrogen molecule (H_2) and the helium atom (He).

5.1 Validation of the discretization basis

We illustrate in this section how to use of the error bars to validate the discretization basis used to solve the Hartree-Fock problem. Error bars are computed for several approximations of the **exact** wavefunction corresponding to several discretization basis and the **exact** (best known) Hartree-Fock energy is seen to fall within the error bars as indicated by the theory. The size of the error bar can be therefore used to to asses the quality of the result and thus to validate the discretisation basis used.

For all the numerical experiments we placed ourselves into the *Restricted (closed) shell Hartree-Fock* (Lewis electron pair) approximation that states that when the number of electrons in a molecule is even, one can group together the electrons 2 by 2; the two electrons in each such pair will share a common spatial wavefunction but will have opposite spin. Within this approximation, for a bi-electronic system as the hydrogen molecule or Helium atom, the search of the electronic wavefunction of the system reduces to the search of a function u of 3 variables such that

$$-\Delta u + Vu + \left(|u|^2 \star \frac{1}{|x|}\right) + \lambda u = 0 \quad in\ \mathbb{R}^3.$$

The space to be discretized is therefore $\mathbb{R}^3$; in fact using classical localization arguments it can be reduced to a brick of $\mathbb{R}^3$ that contains

the nuclei of the system ; in the case of the Helium atom this brick was taken to be a cube centered around the nucleus.

We will present in the following the results obtained for the Helium atom; each axis of a cube centered in the nucleus mentioned above was discretized with the same number of points that varied between 60 and 120 depending on the singularities of the initial solutions considered; precise results were obtained for about 100 points per dimension and corresponding vectors of size $100^3 = 10^6$.

Several initial approximations Φ_i, $i = 2, 3, 4, 5, 6$ of the electronic wavefunction were considered; each correspond to a quantum chemical computation that used specific quantum basis sets denominated as STOnG, $n = 2, 3, 4, 5, 6$; the larger the parameter n, the finer the basis used; in each case the linear problem (30) was solved on the chosen grid as indicated in Remark 12 and then the symmetric part of the error was reconstructed as indicated in previous section. In order to solve (30) an iterative algorithm was employed, the matrix associated to $D^2\mathcal{E}^{\Phi}(\Phi)(\cdot,\cdot)$ (typically $10^6 \times 10^6$) being too large for direct inversion; finally in order to take advantage of the tensor-product-like discretization the computation of convolution products was done by means of fast Fourier transforms.

The figure 1 shows the energy of the initial wavefunction Φ ("Classically computed energy"), the best known approximation of the energy Helium atom, the improved energy obtained as in thm. 2 and then the order ϵ^2 lower bound as described in Thm. 3; agreement with the theoretical results is obtained.

5.2 Validation of the iterative resolution procedure

The numerical resolution of the Hartree-Fock equations involves iterative resolution of eigenvalue problems. The number of iterations necessary is not known in advance and no natural stoping criterion exists. We found therefore important to illustrate how the error bars presented above can be used to validate the number of iterations to be undertaken by the resolution procedure. This time error bars are computed for several approximations of the electronic wavefunction each corresponding to a different number of iterations, **the discretizations basis being kept fixed**. The error bar give in this case lower and upper bounds for the Hartree-Fock energy of the solution of the Hartree-Fock equations on the given discretizations basis. The size of the error bar can be taken as a measure of the improvement still possible if iterations are carried on untill convergence (in the given discrete basis) is reached.

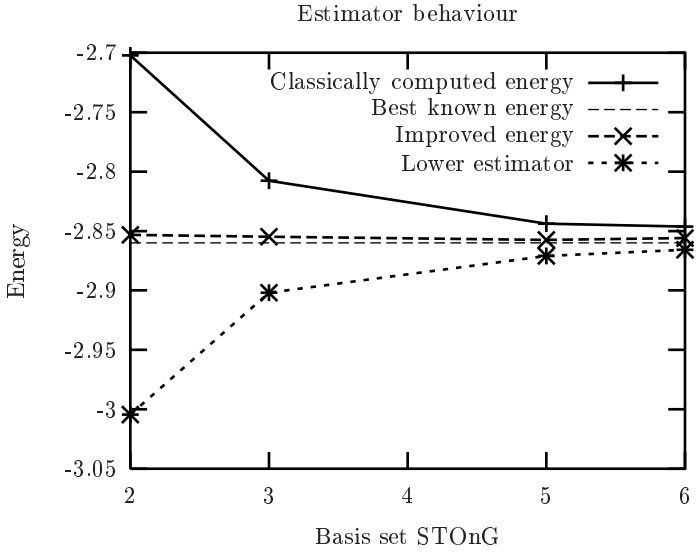

Fig. 1. A posteriori improvements for the energy obtained with the basis sets STOnG.

Motivated by the success of the first series of experiments, this time the molecules considered were larger, as is for instance the case of the carbyne molecule $Cr(CO)_4ClCH$, with 52 electron pairs (104 electrons) ; the model chosen was again the Restricted Hartree Fock model; in this setting the energy to minimize is

$$\mathcal{E}^{HF}(\Phi_1, ..., \Phi_N) = \sum_{i=1}^{N} \int_{\mathbb{R}^3} \left(|\nabla \Phi_i|^2 + V \ |\Phi_i|^2 \right)$$
$$+ \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{\rho_\Phi(x)\rho_\Phi(y)}{|x-y|} dxdy - \frac{1}{2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|\rho_\Phi(x,y)|^2}{|x-y|} dxdy$$

with the same formal definitions (cf. Eq. (3 , 4) for $\rho_\Phi(x)$, $\rho_\Phi(x,y)$). The Euler-Lagrange equations associated to the minimization of $\mathcal{E}^{HF}$ on $\mathcal{H} \cap \mathcal{K}$ are completely similar to (7) (only some multiplicative factors before the last two terms in (5) are changed).

Due to concerns about computation complexity and efficiency and also for realistic verification we have chosen to implement the a posteriori procedure (and the "convergence acceleration" version) in a quantum computational chemistry code named Asterix [7, 19, 23]. As a consequence, the evaluation of the performances of the a posteriori procedure is to be compared with the performances of quantum

chemistry ab initio codes. An introduction to the complexity of the algorithms used is given in the following.

One particularity of computational quantum chemistry codes (especially at the Hartree-Fock level) is the presence of very special Galerkin discretization basis. This basis contains in general functions on $\mathbb{R}^3$ which are centered in the nuclei of the system and are sum of Gaussian type functions; it is beyond the scope of this paper to give a rigorous presentation of the basis involved, let us just say that they all satisfy an important requirement: for any elements h_α, h_β, h_γ and h_δ of the discretization basis, the quantity

$$(\alpha\beta||\gamma\delta) = \iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{h_\alpha(x)h_\beta(x)h_\gamma(y)h_\delta(y)}{|x-y|}dxdy \tag{38}$$

can be computed in $O(1)$ time[3].

Let us denote by n the number of basis functions used when computing the Hartree-Fock energy of a molecule with N electron pairs ($2N$ electrons); in general n is taken to depend linearly on N.

In order to solve the nonlinear eigenvalue equations (7) iterative (also named selfconsistent - SCF) algorithms are used. The most straightforward idea is to start from an initial guess Φ^1 for the wavefunction and then, for any $i \geq 1$, construct the Fock operator $\mathcal{F}^i = \mathcal{F}_{\Phi^i}$ associated to Φ^i, diagonalize $\mathcal{F}^i$ and take its first N eigenfunctions as the next guess Φ^{i+1} for the wavefunction (Roothaan algorithm) ; ideally this fixed point algorithm will converge and the solution will be the solution of equations (7). Numerical reality does not however always validate this choice, we refer to [6] for a mathematical description of the phenomena involved. In order to cure the convergence deficiencies, various other methods have been proposed [6]: the basic level shift method, DIIS,...

During the SCF resolution of the Hartree-Fock equations, the most time consuming part is the construction of the Fock operator $\mathcal{F}_{\Phi^i}$; we will see in the following that this is an $O(N^4)$ operation, one order of magnitude larger than the diagonalization of the Fock operator itself (under assumption that n is linear in N). Let

$$B = \{h_\alpha; \alpha = 1, ..., n\}$$

be a discretization basis and $\Phi = (\sum_{\alpha=1}^n \Phi_{i\alpha}h_\alpha)_{i=1}^N$ be an element in the discretized space $X = (span(B))^N$ and also in $\mathcal{K}$. The matrix of the operators $-\Delta$ and V take $O(N^2)$ time to compute, supposing

[3] Using the fact that the product of two gaussian functions is also a gaussian function, analytical formulas may be provided for the computation of the integral (38).

that finite constant time to compute $\int_{\mathbb{R}^3} \nabla h_\alpha \cdot \nabla h_\beta$ and $\int_{\mathbb{R}^3} V h_\alpha h_\beta$ is needed. The situation is very different for the matrices of the operators $(\rho_\Phi \star \frac{1}{|x|})$ and $\psi \mapsto \int_{\mathbb{R}^3} \frac{\rho_\Phi(x,y)}{|x-y|}\psi(y)dy$. Let us take for instance the last operator. To compute the matrix of this operator it is necessary to compute for all h_β, $h_\gamma \in B$:

$$\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\rho_\Phi(x,y)h_\beta(y)}{|x-y|}dy h_\gamma(x)dx = \sum_{i=1}^{N} \iint_{\mathbb{R}^3\times\mathbb{R}^3} \frac{\sum_{\alpha=1}^n \Phi_{i\alpha}h_\alpha(x)}{|x-y|}$$
$$\cdot \sum_{\delta=1}^{n} \Phi_{i\delta}h_\delta(y)h_\beta(y)h_\gamma(x)dxdy = \sum_{i=1}^{N}\sum_{\alpha=1}^{n}\sum_{\delta=1}^{n} \Phi_{i\alpha}\Phi_{i\delta}(\alpha\gamma||\beta\delta).$$

Even if formally this is a $O(N^5)$ computation (summation over three indices for each of the N^2 required terms), it is easy to see that precomputing in $O(N^3)$ for any $\alpha, \delta = 1, ..., n$: ${D^\Phi}_{\alpha,\delta} = \sum_{i=1}^N \Phi_{i\alpha}\Phi_{i\delta}$ the computation reduces to order N^4 ; unfortunately no further reductions are possible so the matrix of the operator $\psi \mapsto \int_{\mathbb{R}^3} \frac{\rho_\Phi(x,y)}{|x-y|}\psi(y)dy$ is obtained by computing $({D^\Phi}_{\alpha,\delta})_{\alpha,\delta=1}^n$, then obtain in $O(N^4)$ the desired matrix $\left(\sum_{\alpha,\delta=1}^n {D^\Phi}_{\alpha,\delta}(\alpha\gamma||\beta\delta)\right)_{\beta,\gamma=1}^n$. The computational complexity of a SCF Hartree-Fock computation is therefore $N_I \star N^4$ where N_I is the number of iterations required by the SCF method, usually in the range $10 - 50$. We shall apply the bound procedure and the improvement strategy to qualify the (known) solution obtained from the previous iterative procedure far from convergence.

Let us now present the complexity issues related to the computation of the reconstructed error $\hat{W}$. The problem (30) is approximated on a product of N dimensional spaces so the solution will be an $n \times N$ vector (considering the **same** discretization X of $\mathcal{H}$ as the one used to solve the Hartree-Fock problem)[4]; we will denote by P the matrix of the projector from X to $X \cap \Phi^{\perp\!\!\!\perp}$; it is easy to see that P is block diagonal so projecting an element $\Psi = (\sum_{\alpha=1}^n \Psi_{i\alpha}h_\alpha)_{i=1}^n$ of X to $X \cap \Phi^{\perp\!\!\!\perp}$ will be an $O(N^3)$ operation. Let us denote by A_Φ the matrix of the second differential in Φ of the energy with respect to this discretization, and by b_Φ the "vector" corresponding to the

[4] Since only one discretization is used for the entire computation, the bounds thus obtained refer to the energy of the solution of the Hartree-Fock problem on *discrete* space X. When the discretization X is fine enough, one can consider to obtain bounds for the Hartree-Fock energy. In any situation, bounds are usefull e.g. as stopping criteria for the iterative SCF procedure (and eventually to accelerate convergence); then, in order to obtain bounds on the Hartree-Fock energy, correction need to be solved on a grid fine enough to be considered exact as is the case of the computation presented in Fig. 1.

first differential in Φ of the energy, interpreted as an element of the dual X'. The problem (30) has then the following discretization: find $w \in \mathbb{R}^{n\times N}$ such that $w = Pw$ and

$$(P^t A_\Phi P)w + (P^t b_\Phi) = 0. \tag{39}$$

The matrix A_Φ of the linear system (39) is full and impossible to completely invert in practice due to the high computational complexity $O(N^6)$ required. However, using the same argument as above, applying the matrix A_Φ to a vector $v \in \mathbb{R}^{n\times N}$ can be done in $O(N^4)$ operations. The problem (39) is then solved iteratively ; finally let us remark that the total cost of the reconstruction of the symmetric part is an $O(N^3)$ process.

The a posteriori method was tested in the computation of the Hartree-Fock energy of the carbyne $Cr(CO)_4ClCH$ molecule. For each iteration step of the SCF algorithm the order ϵ^4 exact energy estimations were constructed, and also the corresponding lower bounds as described in Thm. 3. The convergence of the SCF method is presented in Fig. 2 and 3. Remark the presence of quadratically convergence periods (iterations 10-50), the presence of "jumps" (55-65) and slow convergence periods (70-90). In order to avoid the last regime, in practice one only uses the SCF algorithm for a small number of iterations 10-40 and then enlarges discretization basis, or tries to empirically optimize other parameters (DIIS).

The results obtained by the a posteriori procedure are presented in the Fig. 4 and 5. For some approximate solution obtained during the SCF iterations, the method described in previous section was applied to improve the energy and obtain a lower bound (initial data corresponding to more than 60 iterations is interpreted as converged due to numerical round-off errors); we do not attach special meaning to the good properties of the reconstructed error for $N = 30$ (cf. Fig. 5). As the results show, the method gives nearly converged results as soon as the initial approximation is as good as the one from the 10^{th} iteration of the SCF procedure.

Remark 14 The number of iterations required to solve the linear system (39) was of the order of 10, which makes this method more efficient than the SCF cycles; for instance finding the improvement from the 10^{th} SCF cycle needs 10 iterations to solve (39) and is as good as the result of the 60^{th} SCF iteration.

Remark 15 Applying the matrix A_Φ to a vector $v \in \mathbb{R}^{n\times N}$ in (39) requires at most $O(N^4)$ operations. The method is however compatible

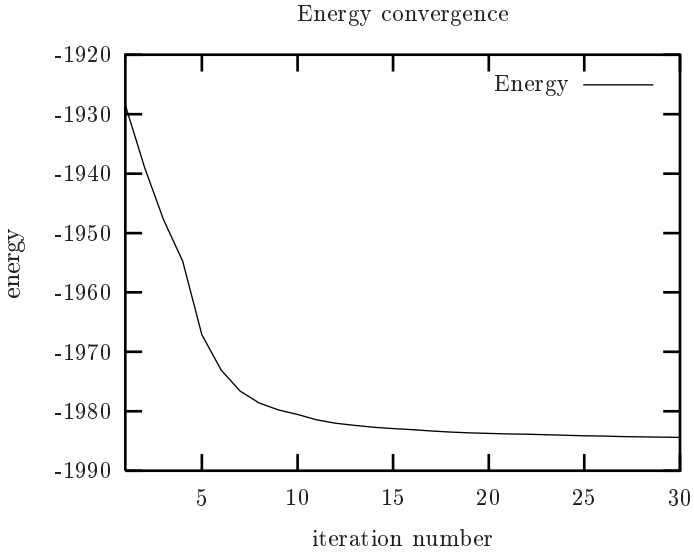

Fig. 2. The convergence of the energy computed by the SCF algorithm in the form used by Chemists. The number of SCF cycles (iterations) ranges between 1 and 30. No a posteriori improvements are made.

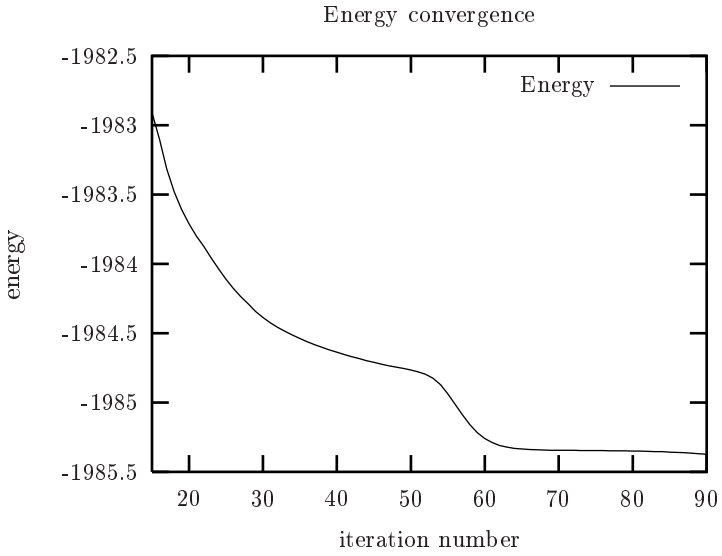

Fig. 3. The convergence of the energy computed by the SCF algorithm in the form used by Chemists. The number of SCF cycles (iterations) ranges between 15 and 90. No a posteriori improvements are made.

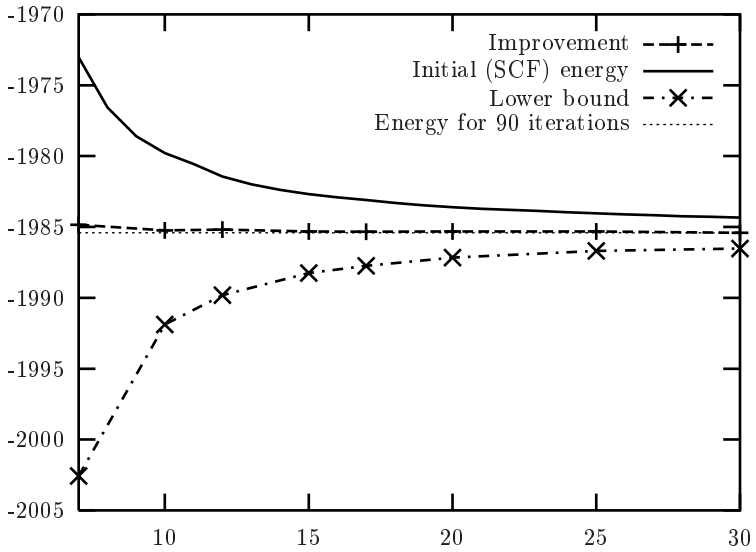

Fig. 4. A posteriori error bounds and improvements are computed for the results of the SCF procedure. In each case we plot the energy of the initial (SCF) approximation, the energy of the wavefunction as computed by the a posteriori improvement procedure and the lower bound as described in Thm. 3. The reference value of the energy is the result of the SCF algorithm after 90 iterations. The initial approximations to improve are the results of the SCF procedure for a number of cycles between 7 and 30.

with the a priori introduction of further localization properties (as domain decomposition methods) of the electronic wavefunction as it is usually the case when more efficient Hartree-Fock computations are searched for [20], which results in the application of the matrix A_Φ being a $O(N^3)$ process (or even less); combining with classical convergence acceleration tools from the linear system solving (preconditioning ...) and with theorem 2, this method can be also seen as another approach towards the design of Hartree-Fock computations of lower algorithmic complexity.

Acknowledgements The authors wish to thank Marie-Madeleine Rohmer and Marc Bénard from "Laboratoire de Chimie Quantique", Strasbourg, France, for providing their code Asterix [7, 19, 23] and for supporting the implementation of the a posteriori method in Asterix.

References

1. I. Babuška "A posteriori error estimation for the finite element method", Internat. J. Numer. Methods Engrg. **12** (1978), 1597-1615

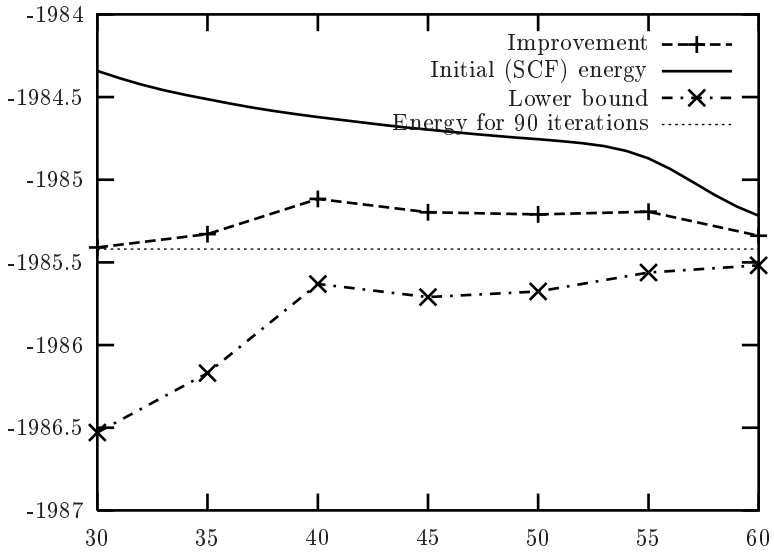

Fig. 5. See Fig. 4 for details. The initial approximations to improve are the results of the SCF procedure for a number of cycles between 30 and 60.

2. I. Babuška and C. Schwab "A posteriori error estimation for hierarchic models of elliptic boundary value problems on thin domains" SIAM J. Numer. Anal. Vol **33** (1996), No.1, pp 241-246
3. G. B. Bacskay, "A Quadratically Convergent Hartree-Fock (QC-SCF) Method. Application to Closed Systems," Chem. Phys. 61, 385-404 (1981).
4. C. Bernardi, B. Metivet, "Indicateurs d'erreur pour l'equation de la chaleur. (Error indicators for the heat equation)" Rev. Eur. Elem. Finis 9, No.4, 425-438 (2000).
5. Eric Cances, Ecole Nationale des Ponts et Chaussées, Marne la Vallée, France, private communication, january 2001.
6. E.Cancès and C.LeBris, "On the convergence of SCF algorithms for the Hartree-Fock equations", M2AN, Vol. 34, No. 4, 2000, pp. 749-774
7. R. Ernenwein, M-M. Rohmer and M. Bénard, "A program system for ab initio MO calculations on vector and parallel processing machines, I. Evaluation of integrals", Comput. Phys. Comm. 58(1990), p.305-328
8. S. Goedecker "Linear scaling electronic structure methods", Rev. Mod. Phys vol. 71, No.4, July 1999, p 1085-1123
9. P.Ladevèze and D.Leguillon "Error estimate procedure in the finite element method and applications" SIAM J.Numer Anal. Vol 20,1991, no 5, 485-504
10. E.H.Lieb and B.Simon " The Hartree-Fock theory for Coulomb systems" Commun. Math. Phys. 53, 185-194 (1977)
11. P.L.Lions "Solutions of Hartree-Fock Equations for Coulomb systems" Commun. Math. Phys. 109, 33-97(1987)
12. Y. Maday, A.T. Patera "Numerical analysis of a posteriori finite element bounds for linear-functional outputs", Math. Models Methods Appl. Sci. 10 (2000), no. 5, 785–799.

13. Y. Maday, A. T. Patera and J. Peraire, "A general formulation for a posteriori bounds for output functionals of partial differential equations; application to the eigenvalue problem", Comptes Rendus de l'Académie des Sciences - Serie I - Mathematiques (328) 9 (1999) pp. 823-828
14. J.T.Oden and Y.Feng, "Local and pollution error estimation for finite element approximations of elliptic boundary value problems" J.Comput. Appl. Math, **74**, 245-293 (1996)
15. M. Paraschivoiu, A.T. Patera, "A hierarchical duality approach to bounds for the outputs of partial differential equations", Computer Methods in Applied Mechanics and Engineering 158 (3-4) (1998) pp. 389-407.
16. M. Paraschivoiu, J. Peraire, A.T. Patera, "A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations", Computer Methods in Applied Mechanics and Engineering 150 (1-4) (1997) pp. 289-312.
17. J. Peraire, A.T. Patera, "Asymptotic a posteriori finite element bounds for the outputs of noncoercive problems: the Helmholtz and Burgers equations", Computer Methods in Applied Mechanics and Engineering 171 (1-2) (1999) pp. 77-86.
18. J. Pousin et J. Rappaz, "Consistency, stability, a priori and a posteriori errors for Petrov-Galerkin methods applied to nonlinear problems" (Report), EPFL, Lausanne, 1992
19. M-M. Rohmer, J. Demuynck, M. Bénard, R. Wiest, C. Bachmann, C. Henriet and R. Ernenwein "A program system for ab initio MO calculations on vector and parallel processing machines, II. SCF closed-shell and open shell iterations", Comput. Phys. Comm. 60(1990), p.127-144
20. M. C. Strain, G. E. Scuseria and M. J. Frisch, "Achieving linear scaling for the electronic quantum Coulomb problem." Science 271, 51 (1996).
21. R.Verfürth "A Review of A Posteriori Error Estimates and Adaptative Mesh-Refinement Techniques", Wiley-Teubner 1997
22. R.Verfürth "A Posteriori Error Estimates For Non-Linear Problems. Finite Element Discretisations of Elliptic Equations" Math. of Comp. **62**, 206(1994), pp 445-475
23. R. Wiest, J. Demuynck, M. Bénard, M-M. Rohmer and R. Ernenwein "A program system for ab initio MO calculations on vector and parallel processing machines, III. Integral reordering and four-index transformation", Comput. Phys. Comm. 62(1991), p.107-124