

Solving Hartree-Fock systems with global optimization methods

C. LAVOR^{1(a)}, L. LIBERTI^{2(b)}, N. MACULAN^{3(c)} and M. A. C. NASCIMENTO^{4(d)}

¹ *Department of Applied Mathematics (IMECC-UNICAMP), State University of Campinas - CP 6065, 13081-970, Campinas-SP, Brazil*

² *LIX, École Polytechnique - F-91128 Palaiseau, France*

³ *COPPE, Universidade Federal do Rio de Janeiro, UFRJ - CP 68511, Rio de Janeiro - RJ, 21945-970, Brazil*

⁴ *Departamento de Físico-Química, Instituto de Química, Universidade Federal do Rio de Janeiro, UFRJ, Rio de Janeiro - RJ, 21949-970, Brazil*

received 21 September 2006; accepted in final form 11 January 2007

published online 23 February 2007

PACS 02.60.Pn – Numerical optimization

PACS 31.15.Ar – *Ab initio* calculations

PACS 71.15.Dx – Computational methodology (Brillouin zone sampling, iterative diagonalization, pseudopotential construction)

Abstract – The Hartree-Fock equations describe atomic and molecular electronic wave functions, based on the minimization of a functional of the energy. This can be formulated as a constrained global optimization problem involving nonconvex polynomials exhibiting many local minima. The traditional method of solving the Hartree-Fock problem does not provide a guarantee of global optimality and is very sensitive to the initial starting point. In this paper we show how to use a deterministic global optimization method to solve Hartree-Fock systems. The validity of the proposed approach was established by successfully computing the ground-state of the He and Be atoms.

Copyright © EPLA, 2007

Introduction. – The quantum behaviour of atoms and molecules, in the absence of relativistic effects, is determined by the time-independent Schrödinger equation:

$$H\Psi_n = E\Psi_n, \quad (1)$$

where H , the Hamiltonian operator of the system, represents the total energy (kinetic + potential) of all the particles of the system.

Analytical solutions for this equation are only possible for very simple systems. Hence, for the majority of the problems of interest, one has to rely on some approximate model. In the Hartree-Fock model, the electrons in atoms and molecules move independently of each other, the motion of each one of the electrons being determined by the attractive electrostatic potential of the nuclei and by a repulsive average field due to all the other electrons of the system. In this model, the approximate solutions Φ_n of eq. (1) are anti-symmetrized products of one-electron wave functions $\{\varphi_i\}$ (also called orbitals), which

are solutions of the Hartree-Fock (HF) equations for the system under study [1]. This model gives rise to a set of coupled integro-differential equations which can only be solved numerically. Alternatively, each orbital φ_i can be expanded in a complete basis set $\{\chi_s\}_{s=1}^{\infty}$. In order to transform the HF equations into a less cumbersome algebraic problem, we only consider a non-orthogonal finite subset $\{\chi_s | s \leq b\}$ of the basis, and we use it to approximate the orbitals. The larger we choose b , the better the approximation is likely to become.

The optimization problem considered in this paper arises because we need to find a set of coefficients c_{si} , for $s = 1, \dots, b$ and $i = 1, \dots, n$, such that for all $i \leq n$ the function

$$\bar{\varphi}_i = \sum_{s=1}^b c_{si} \chi_s \quad (2)$$

is a good approximation of the i -th spatial orbital φ_i . A further requirement on the approximating set $\{\bar{\varphi}_i\}$ is that it must be an orthogonal set. While not a necessary condition, orthogonality is always imposed in the Hartree-Fock method because the resulting equations are much easier to solve in an orthogonal basis of atomic or molecular orbitals [2]. The method most usually applied

^(a)E-mail: clavor@ime.unicamp.br

^(b)E-mail: liberti@lix.polytechnique.fr

^(c)E-mail: maculan@cos.ufrj.br

^(d)E-mail: chaer@iq.ufrj.br

to the Hartree-Fock equations iteratively solves a set of linear equations to find the coefficients c_{si} . This method, however, has two main limitations: (a) there is no guarantee that the set of coefficients c_{si} found by the method is a globally optimal set; (b) it depends on an initial solution being available (starting guess); (c) the occupation number of all orbitals must be provided (electronic configuration).

We define the Hartree-Fock Problem (HFP) as the problem of finding a set of coefficients c_{si} such that the $\bar{\varphi}_i$ are the best possible approximations of the spatial orbitals. The objective function (quality of the approximation) is given by the energy function E associated with the approximating set $\{\bar{\varphi}_i\}$, which is guaranteed to be an upper bound to the energy function associated with the spatial orbitals [3]. The set $\{\bar{\varphi}_i\}$ is required to be orthonormal for the reasons already mentioned. Furthermore, once each member of the set $\{\bar{\varphi}_i\}$ is expanded in a complete orthonormal basis $\{\chi_s\}$, a normalization condition must be imposed on the expansion coefficients in order to preserve the probabilistic interpretation of the wave function. Thus, we need to minimize the energy function E subject to orthonormality and normalization constraints. The decision variables of this mathematical programming problem are the coefficients c_{si} . For the orthonormal basis sets $\{\bar{\varphi}_i\}$ and $\{\chi_s\}$, the problem can be expressed as follows:

$$\begin{cases} \min_c & E(c) \\ \text{subject to (s.t.)} & \langle \bar{\varphi}_i | \bar{\varphi}_j \rangle = \delta_{ij} \quad \forall i \leq j \leq n \\ & \sum_{r=1}^b c_{ri}^2 = 1 \quad \forall i \leq n \\ & c^L \leq c \leq c^U \end{cases} \quad (3)$$

where δ_{ij} is the Kronecker delta-function, which is equal to 1, if $i = j$ and 0, otherwise. Problem (3) is a nonconvex, multi-extremal, polynomially constrained, polynomial programming problem, and falls therefore in the realm of Global Optimization (GO). We solve it by applying the spatial Branch-and-Bound (sBB) technique, a well-known deterministic GO method which yields an ε -guaranteed global optimum (for a given $\varepsilon > 0$) and does not need an initial solution, *i.e.* a set of starting guess coefficients. For validation purposes, this method has been applied to two instances of the HFP, namely to the helium and berillium atoms, with considerable success as regarding the CPU performance. In spite of the fact that He and Be are closed-shell systems, the sBB technique can be easily extended to open-shell systems (ROHF) and also to more sophisticated multiconfigurational wave functions, as will be discussed in future publications.

The rest of this paper is organized as follows. Next section summarizes the sBB algorithm for GO. The third section presents the mathematical programming formulation of the HFP. The fourth section describes the method used to compute the lower bound to the objective function value. The fifth section discusses the computational experience on the He and Be atoms. Last section concludes the paper.

Global Optimization. – A mathematical programming problem is formulated as follows:

$$\begin{cases} \min_x & f(x) \\ \text{s.t.} & g(x) \leq 0, \end{cases} \quad (4)$$

where $x \in \mathbb{R}^n$ are the *decision variables* and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is the *objective function* to be minimized subject to a set of constraints $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ which may also include variable ranges or integrality constraints on the variables. Global Optimization is concerned with the solution of problems (4) where f, g are non-convex nonlinear forms. A problem where f, g are nonlinear is known as a Non Linear Programming problem (NLP); if some integrality constraints are present on the variable bounds the problem is known as a Mixed-Integer NLP (MINLP).

GO methods may be deterministic or stochastic according as to whether they provide a guarantee of optimality or they employ random search techniques. The most commonly employed deterministic GO method is an extension, called spatial Branch-and-Bound (sBB) [4–6], of the well-known Branch-and-Bound algorithm for implicit binary enumeration [8] to continuous spaces. In sBB, branching occurs by partitioning the continuous variable range in two or more sub-ranges. The recursive application of branching gives rise to a search tree. A node is *fathomed* (*i.e.* no further branching occurs on the node) either because the global minimum relative to the node has been found, or because the global minimum relative to the node cannot be better than the overall best solution found so far (the *incumbent*). In order to test these two conditions at each node, we compute a lower bound and an upper bound to the objective function value of the node's problem restriction. The first condition is verified if these bounds differ by less than a pre-specified $\varepsilon > 0$ tolerance, and the second if the lower bound for the node is higher than the incumbent.

Because Branch-and-Bound algorithms have exponential worst-case complexity, the best way to run them efficiently is to provide a tight lower bound. In the specific case of the HFP, the lower bound is provided by solving a particular linear relaxation of the problem strengthened by several *valid inequalities* based on the quadratic structure of the problem and on the orthogonality constraints.

Mathematical programming formulation of the problem. – The expression for the Hartree-Fock electronic energy E of a $2n$ -electron molecule with closed shells is given by [3] as

$$E = 2 \sum_{i=1}^n H_{ii}^{core} + \sum_{i=1}^n \sum_{j=1}^n (2J_{ij} - K_{ij}) + V_{NN}, \quad (5)$$

where H_{ii}^{core} contains the one-electron integrals, J_{ij} is the Coulomb integral, K_{ij} is the exchange integral, V_{NN} is the nuclear repulsion term. The spatial orbitals $\{\varphi_i | i \leq n\}$ are expanded as linear combinations of a finite set of one-electron basis functions as per eq. (2). However,

the standard self-consistent procedure to solve the HF equations uses a non-orthogonal basis set $\{\chi_s\}$. Thus, in order to compare our results with the ones obtained by the standard procedure, the sBB technique will be also formulated in a non-orthogonal basis set $\{\chi_s\}$. The HF equations [3] are a set of b equations in the variables c_{si} :

$$\sum_{s=1}^b c_{si} (F_{rs} - \varepsilon_i S_{rs}) = 0, \quad (6)$$

where $F_{rs} = \langle \chi_r | \hat{F} | \chi_s \rangle$, $S_{rs} = \langle \chi_r | \chi_s \rangle$, ε_i is the orbital energy, and \hat{F} is the Fock operator [3]. This is a nonlinear system, since the \hat{F} operator depends on the orbitals $\{\varphi_i\}$, which in turn depend on the variables c_{si} .

It is possible to obtain an expression for F_{rs} in terms of the coefficients $\{c_{si}\}$ and a set of suitable integrals over the basis functions $\{\chi_s\}$ [3]:

$$F_{rs} = H_{rs}^{core} + \sum_{t=1}^b \sum_{u=1}^b \sum_{i=1}^n c_{ti}^* c_{ui} [2(rs|tu) - (ru|ts)], \quad (7)$$

where $(rs|tu)$ and $(ru|ts)$ stand for the Coulomb and exchange integrals between pairs of electrons. While the Coulomb integrals represent the quantum-mechanical equivalent of the classical Coulomb interaction between two charged particles, the exchange integrals are purely quantum entities, resulting from the fact that the total wave function for any multi-electronic system must be anti-symmetric (Pauli principle). Also, the wave functions representing the behaviour of atoms and molecules can be real or complex. However, since any observable (dynamic variable) must be real, the expectation value of the corresponding operator, O , must be taken as $\langle O \rangle = \int \phi^* O \phi dv$, which is the reason why the complex product $c_{ti}^* c_{ui}$ appears in eq. (7).

We write $H(r, s) = H_{rs}^{core}$ and $X(r, s, t, u) = (rs|tu)$, where $r = 1, \dots, b$, $s = 1, \dots, b$, $t = 1, \dots, b$, and $u = 1, \dots, b$. It is possible to rewrite eq. (5) in the form

$$E = \sum_{r=1}^b \sum_{s=1}^b \sum_{i=1}^n (c_{ri}^* c_{si} (F_{rs} + H(r, s))) + V_{NN}. \quad (8)$$

Finally, using (7) and (8), we get

$$\begin{aligned} E = & \frac{1}{2} \sum_{r=1}^b \sum_{s=1}^b \sum_{t=1}^b \sum_{u=1}^b \left(P(r, s) P(t, u) \right. \\ & \times \left. \left(X(r, s, t, u) - \frac{1}{2} X(r, u, t, s) \right) \right) \\ & + \sum_{r=1}^b \sum_{s=1}^b (P(r, s) H(r, s)) + V_{NN}, \end{aligned} \quad (9)$$

where

$$P(j, k) = 2 \sum_{i=1}^n c_{ji}^* c_{ki} \quad (j = 1, \dots, b \text{ and } k = 1, \dots, b).$$

Note that expression (9) is actually a function of the coefficients c_{si} , since the integrals $H(r, s)$ and $X(r, s, t, u)$, as well as the value of V_{NN} , can be calculated once the basis $\{\chi_s\}$ and the molecular geometry are defined.

As has been mentioned above, the decision variables of the HFP are the coefficients c_{si} used in the basis expansion of the spatial orbitals. To further simplify the notation, we shall write the numerical problem parameters as

$$\begin{aligned} \alpha_{rs}^{tu} &= X(r, s, t, u) - \frac{1}{2} X(r, u, t, s), \\ \beta_{rs} &= H(r, s), \\ \gamma &= V_{NN}. \end{aligned}$$

After simple term rearrangement, the objective function of the problem becomes

$$\begin{aligned} E(c) = & 2 \sum_{r,s=1}^b \left(\sum_{i=1}^n c_{ri} c_{si} \right) \\ & \times \left(\sum_{t,u=1}^b \alpha_{rs}^{tu} \left(\sum_{i=1}^n c_{ti} c_{ui} \right) + \beta_{rs} \right) + \gamma. \end{aligned} \quad (10)$$

The orthogonality constraints are $\langle \bar{\varphi}_i | \bar{\varphi}_j \rangle = \delta_{ij}$ for all $i \leq j \leq n$. Substituting $\bar{\varphi}_i = \sum_{r \leq b} c_{ri} \chi_r$ for all $i \leq n$ and $\langle \chi_r | \chi_s \rangle = S_{rs}$ for all $r, s \leq b$ we obtain

$$\sum_{r,s \leq b} S_{rs} c_{ri} c_{sj} = \delta_{ij} \quad \forall i \leq j \leq n, \quad (11)$$

where normally $S_{rr} = 1$ for all $r \leq b$. The variable bounds

$$c^L \leq c \leq c^U \quad (12)$$

depend on the instance. The HFP can be succinctly summarized as follows:

$$\min\{E(c) \mid \text{s.t. (11)-(12)}\}, \quad (13)$$

as the minimization of a quartic objective function subject to quadratic constraints and variable bounds.

Computing the lower bound. – Recall that the most important step in the efficient running of the sBB algorithm is the computation of the lower bound at each node of the sBB search tree. In this section we shall discuss a method which produces reasonably tight bounds for the HFP. At any given sBB node, some of the decision variables are restricted to lie in subranges given by the recursive branching rules. We consider the restriction of the original problem (13) to the node's variable ranges. We mention in passing that a reasonably good upper bound can be easily provided for by finding any local minimum of (13) restricted to the node (in practice, this is done by using a constrained nonlinear local solver such as SNOPT [9] on the problem formulation). The lower bound is calculated by finding a linear relaxation of the problem restricted to the node and solving it with Linear Programming (LP) tools (for example the

simplex method [10] implemented in CPLEX [11]). The linear relaxation is built in three steps: (a) reformulation, (b) relaxation, (c) tightening.

Step (a) reformulates the problem to a standard form [5] consisting of a linear objective function subject to linear constraints and a set of quadratic constraints. By introducing new variable sets y, w with the following constraints:

$$y_{rs}^{ij} = c_{ri}c_{sj}, \quad \forall r, s \leq b, \quad i, j \leq n, \quad (14)$$

$$w_{rstu}^{ij} = y_{rs}^{ii}y_{tu}^{jj}, \quad \forall r, s, t, u \leq b, \quad i, j \leq n, \quad (15)$$

we can replace all the quartic and quadratic terms in the problem, obtaining

$$\left\{ \begin{array}{l} \min_y \quad 2 \sum_{r,s \leq b} \left(\sum_{t,u \leq b} \alpha_{rs}^{tu} \sum_{i,j \leq n} w_{rstu}^{ij} + \beta_{rs} \sum_{i \leq n} y_{rs}^{ii} \right) + \gamma \\ \text{s.t.} \quad \sum_{r,s \leq b} S_{rs} y_{rs}^{ij} = \delta_{ij}, \quad \forall i \leq j \leq n \\ \sum_{r \leq b} y_{rr}^{ii} = 1, \quad \forall i \leq n \\ \begin{array}{l} (14)-(15) \\ c^L \leq c \leq c^U \\ y^L \leq y \leq y^U \\ w^L \leq w \leq w^U, \end{array} \end{array} \right. \quad (16)$$

where the bounds y^L, y^U on y and w^L, w^U on w are obtained through simple interval arithmetics using the bounds on c and the quadratic relations (14)-(15).

Step (b) constructs a relaxation of (16) by replacing (14)-(15) by their convex envelopes. For simple constraints of the form $z = xt$ where $x^L \leq x \leq x^U$ and $t^L \leq t \leq t^U$ (such as (14)-(15)) the convex envelope is given by [12,13]:

$$z \leq \min\{x^U t + t^L x - x^U t^L, x^L t + t^U x - x^L t^U\} \quad (17)$$

$$z \geq \max\{x^L t + t^L x - x^L t^L, x^U t + t^U x - x^U t^U\}, \quad (18)$$

which defines an enveloping tetrahedron around the points $(x^L, t^L), (x^L, t^U), (x^U, t^L), (x^U, t^U)$. Purely quadratic constraints of the form $z = x^2$ are relaxed by the secant and the tangents of the parabola at $(x^L, (x^L)^2), (x^U, (x^U)^2)$. This relaxation yields a linear problem whose optimal objective function value \bar{f} is a lower bound to the globally optimal objective function value f^* of (16), which is the same as that of (13). Since (17)-(18) vary as a function of the variable bounds, \bar{f} depends on c^L, c^U , which means that the lower bound depends on the current sBB search tree node, as desired. In practice, it turns out that \bar{f} is not a very tight bound. This is mostly due to the fact that the envelope (17)-(18) is generally not very close to the original surface (14)-(15). We try to improve this situation by adding some valid inequalities in the next step.

Step (c) generates valid constraints in order to tighten the linear relaxation obtained in steps (a) and (b).

Table I: Computational results for the He and Be atoms.

Atom	sBB	VNS	SobolOpt
He	0.26 s	0.116 s	0.14 s
Be	10 s	0.3 s	14 s

Recalling that $S_{rr} = 1$ for all $r \leq b$, we first remark that when $i = j$, constraints (11) are

$$\sum_{r=1}^b c_{ri}^2 + 2 \sum_{r < s} S_{rs} c_{ri} c_{si} = 1, \quad \forall i \leq n, \quad (19)$$

which can be written in terms of the y variables as

$$\sum_{r=1}^b y_{rr}^{ii} + 2 \sum_{r < s} S_{rs} y_{rs}^{ii} = 1, \quad \forall i \leq n. \quad (20)$$

Secondly, we multiply (20) by problem variables y_{tu}^{jj} for all $t, u \leq b, j \leq n$, obtaining

$$\sum_{r=1}^b y_{rr}^{ii} y_{tu}^{jj} + 2 \sum_{r < s} S_{rs} y_{rs}^{ii} y_{tu}^{jj} = y_{tu}^{jj}, \quad \forall t, u \leq b, \quad i, j \leq n.$$

We can now replace the quadratic products in the y variables using the w variables as per (15), obtaining a linear relation between the y and the w variables:

$$\sum_{r=1}^b w_{rrtu}^{ij} + 2 \sum_{r < s} S_{rs} w_{rstu}^{ij} = y_{tu}^{jj}, \quad \forall t, u \leq b, \quad i, j \leq n. \quad (21)$$

In the linear relaxation of (16) the y and w variables are only related through inequality constraints of type (17)-(18). Constraints (21) tighten the relaxation because in general a feasible region defined by equations has a smaller measure than one defined by inequalities ([14], theorem 4.1). Constraints (21) are called *reduced RLT constraints* [14–16] and form a subclass of the Reformulation-Linearization Technique (RLT) constraints described in [17].

Computational results. – We performed some preliminary computational tests on the ground state of the helium (He) and berillium (Be) atoms. For the He atom we considered an uncontracted basis of two s -type Gaussian functions and for the Be atom we used the STO-3G basis set, as detailed in [18]. The globally optimal solutions are $E_{\text{He}}^* = -2.7471h$ and $E_{\text{Be}}^* = -14.3519h$, and both were confirmed by the sBB algorithm. These results are in perfect agreement with the ones obtained through the standard self-consistent procedure to the HF problem. We performed some computational comparisons in terms of user CPU time (reported in table I) by solving the problem instances with two other global optimization solvers, namely Variable Neighbourhood Search (VNS) [19,20] and SobolOpt [21] (which are stochastic and do not provide an optimality guarantee).

All computational results have been obtained within the *ooOPS* optimization software framework [6,22] running Linux on an Intel Pentium III 850 MHz with 384 MB RAM.

Conclusion. – The usual way to solve Hartree-Fock equations has two limitations: (a) it provides a solution which is not guaranteed to be the optimal one, (b) it depends heavily on an initial solution being provided and (c) the occupation number of all orbitals must be provided (electronic configuration). We formulate this problem as a nonconvex optimization problem, which we solve using a spatial Branch-and-Bound algorithm for Global Optimization. The crucial step, *i.e.* the determination of the lower bound at each search tree node, relies on a convex relaxation which is further tightened by a suitable class of valid constraints. This approach overcomes both limitations of the usual method. The computational results presented in this paper refer to rather small test cases, but are nonetheless very promising. Computational work on larger cases is ongoing.

The authors are thankful to FAPERJ, FAPESP and CNPq for their support. One of the authors (LL) is thankful to Prof. D. KROB for financial support.

REFERENCES

- [1] Roothan C. C. J., *Rev. Mod. Phys.*, **23** (1951) 69.
- [2] Löwdin P. O., *J. Chem. Phys.*, **18** (1950) 365.
- [3] Levine I. N., *Quantum Chemistry*, second edition (Prentice-Hall, Upper Saddle River, New Jersey) 2000.
- [4] Adjiman C. S., Dallwig S., Floudas C. A. and Neumaier A., *Comput. Chem. Eng.*, **22** (1998) 1137.
- [5] Smith E. M. B. and Pantelides C. C., *Comput. Chem. Eng.*, **23** (1999) 457.
- [6] Liberti L., *Writing global optimization software*, in [7], pp. 211-262.
- [7] Liberti L. and Maculan N. (Editors), *Global Optimization: From Theory to Implementation* (Springer, Berlin) 2006.
- [8] Aho A. V., Hopcroft J. E. and Ullman J. D., *Data Structures and Algorithms* (Addison-Wesley, Reading, MA) 1983.
- [9] Gill P. E., *User's Guide for SNOPT 5.3*, Systems Optimization Laboratory, Department of EESOR, Stanford University, California, February 1999.
- [10] Dantzig G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, NJ) 1963.
- [11] ILOG, *ILOG CPLEX 8.0 User's Manual*, ILOG S.A., Gentilly, France, 2002.
- [12] McCormick G. P., *Math. Program.*, **10** (1976) 146.
- [13] Al-Khayyal F. A. and Falk J. E., *Math. Oper. Res.*, **8** (1983) 273.
- [14] Liberti L. and Pantelides C. C., *J. Global Opt.*, **36** (2006) 161.
- [15] Liberti L., *Int. Transact. Oper. Res.*, **11** (2004) 34.
- [16] Liberti L., *J. Global Opt.*, **33** (2005) 157.
- [17] Serali H. D. and Alameddine A., *J. Global Opt.*, **2** (1992) 379.
- [18] Lavor C., Cardozo T. M. and Nascimento M. A. C., *Inter. J. Quantum Chem.*, **103** (2005) 500.
- [19] Mladenović N., Petrović J., Kovačević-Vujčić V. and Čangalović M., *Eur. J. Oper. Res.*, **151** (2003) 389.
- [20] Liberti L. and Dražić M., *Variable neighbourhood search for the global optimization of constrained NLPs*, in *Proceedings of GO Workshop, Almeria, Spain, 2005*.
- [21] Kucherenko S. and Sytsko Yu., *Comput. Opt. Appl.*, **30** (2004) 297.
- [22] Liberti L., Tsiakis P., Keeping B. and Pantelides C. C., *ooOPS*, Centre for Process Systems Engineering, Chemical Engineering Department, Imperial College, London, UK, 2001.