A mathematical programming formulation for the Hartree-Fock Problem on open-shell systems

Leo Liberti · **Carlile Lavor** · **Nelson Maculan**

Abstract The solutions of the time-independent Schrödinger equation provide a quantum description of the stationary state of electrons in atoms and molecules. The Hartree-Fock problem consists in expressing these solutions by means of finite dimensional approximations thereof. These are themselves linear combinations of an existing linearly independent set; best approximations are obtained when a certain energy function is minimized. In [12] we proposed a new Mathematical Programming (MP) approach which enhanced the likelihood of attaining globally optimal approximations, limited to closed-shell atomic systems. In this paper, we discuss an extension to open-shell systems: this is nontrivial as it requires the expression of a rank constraint within an MP formulation. We achieve this by explicitly modelling eigenvalues and requiring them to be nonzero. Although our approach might not necessarily scale well, we show it works on two open-shell systems (lithium and boron).

Keywords quantum chemistry, mathematical programming, linear independence, rank constraint.

1 Introduction

The wave function of a quantum many-body system of electrons in atoms and molecules satisfies the time-independent Schrödinger equation:

$$H\Phi = E\Phi. \tag{1}$$

L. Liberti
CNRS LIX, École Polytechnique, F-91128 Palaiseau, France
E-mail: `liberti@lix.polytechnique.fr`

C. Lavor
Department of Applied Mathematics (IMECC-UNICAMP), University of Campinas, 13081-970, Campinas-SP, Brazil
E-mail: `clavor@ime.unicamp.br`

N. Maculan
Universidade Federal do Rio de Janeiro, COPPE – UFRJ, CP 68511, Rio de Janeiro – RJ, 21945-970, Brazil
E-mail: `maculan@cos.ufrj.br`

This equation is the projection on the position/momentum variables of the time-dependent Schrödinger equation [24], which describes the changes over time of a quantum physical system. The time-dependent Schrödinger equation describes the wave nature of particles and explains phenomena such as quantum jumps, quantum state superposition, and probabilistic interpretations of particle position and momentum [21]. The time-independent version given in Eq. (1) describes the standing waves corresponding to systems in some form of equilibrium. H is a differential operator which represents the total energy of the system. Solutions to Eq. (1) are usually impossible to derive analytically (aside from the very simplest cases), so approximate numerical solutions are usually sought instead.

The Hartree-Fock method is used to approximate Φ by expressing it as a linear combination of finitely many basis functions. We propose a Mathematical Programming (MP) formulation to decide the value of these coefficients. More specifically, the approximations of Φ are anti-symmetrized products of a set $\{\bar{\varphi}_i \mid i \leq n\}$ of one-electron wave functions called *spatial orbitals*, each of which is approximated by a linear combination of a finite subset $B \subseteq \mathbb{N}$ of a complete basis set $\{\chi_r\}_{r=1}^{\infty}$ of the Hilbert space. The larger $|B|$ is, the better the approximation is likely to become. See [18, §1] for more information.

The Hartree-Fock problem (HFP) was defined in [12]. To summarize, we want to find a set of coefficients c_{ri}, for $r \in B$ and $i \in N = \{1, \ldots, n\} \subseteq B = \{1, \ldots, b\}$, such that:

$$\forall i \in N \quad \varphi_i = \sum_{r \in B} c_{ri} \chi_r \qquad (2)$$

are good approximations of the spatial orbitals $\bar{\varphi}_i$. The quality of the approximation is inversely proportional to the energy function E associated with the approximating set $\{\varphi_i\}$, which is guaranteed to be an upper bound to the energy function associated with the spatial orbitals [13]. The most common procedure used to solve the HFP is the Self-Consistent Field (SCF) methodology [27]. This is essentially an iterative optimization algorithm that takes an initial solution as input and improves it until a local optimum is reached. As such, its main limitations are that: (a) it needs a good starting point, and (b) it may fail to find the global optimum.

The expression for the energy function E varies according to whether the considered system is open- or closed-shell. Unlike a closed-shell system, an open-shell system has a molecular orbital with a valence shell which is not completely filled with electrons, or that has not given all of its valence electrons through chemical bonds (see `en.wikipedia.org/wiki/Open_shell`).

The HFP was first cast as a nonconvex Nonlinear Program (NLP) in [11]: in that paper, an interval Branch-and-Bound (iBB) method was proposed to find $\{\varphi_i\}$ for the He (helium) and Be (beryllium) closed-shell atomic systems. In [12, 18], a different type of Branch-and-Bound (BB) algorithm, called spatial BB (sBB), was tested on the same atomic systems He and Be. Both iBB and sBB provide an optimality guarantee to within a given $\varepsilon > 0$ tolerance, but sBB turns out to be more computationally efficient than the iBB on the HFP. More guaranteed results (H_2, LiH and H_3^+) were obtained in [9, 10] using a specially adjusted version of the new, open-source sBB solver COUENNE [1]. Some work in the same direction was also carried out in [28]. All the MP-based methods reviewed above refer to closed-shell atomic systems: the main properties of such systems are (a) they can be approximated by expressing them as linear combinations of a finite-dimensional function basis (b) they minimize a certain energy function and (c) the set $\{\varphi_i\}$ turns out to be an orthonormal set (see [12] and [19, §3.5]). These properties are exploited in MP formulations as summarized in Sect. 2 below.

The present paper proposes an MP-based method to address the extension of the HFP to open-shell systems. In such systems, the orthonormality of $\{\varphi_i\}$ is not a physical requirement: in fact, all that is required is linear independence [19]. Orthonormality is usually imposed to simplify the SCF calculations [8]. A consequence of this fact is that the SCF methodology is even more likely to find locally optimal solutions which fail to be global. Linear independence restrictions amount to a rank constraint on a matrix of decision variables, and is very difficult to express in MP formulations. A general modelling methodology was proposed in [15]. In this paper we derive a new MP formulation for the HFP on open-shell systems (Sect. 3), and solve it with a combination of reformulations [14] and Global Optimization (GO) methods (Sect. 4) on the B (boron) and Li (lithium) atomic systems (Sect. 5).

2 Formulation for closed-shell systems

In this section we recall from [12,18] the MP formulation for the HFP on closed-shell systems, which is based on orthogonality constraints on the orbitals. We employ an approximating basis set $X = \{\chi_r \mid r \in B\}$. We consider three sets of parameters α, β, γ related to the energy of the considered system: α depends on four indices $r, s, t, u \in B$, β depends on two indices $r, s \in B$ and γ is a scalar. Since we have to decide how the approximated orbitals φ_i depend on the approximating basis X, by (2) we have to find an assignment of feasible values for the coefficients c_{ri} (for $r \in B, i \in N$). In other words, c_{ri} are the decision variables of our MP formulation. The molecular energy is given in [11,12]:

$$E^{\text{closed}}(c) = 2 \sum_{r,s \in B} \left(\sum_{i \in N} c_{ri} c_{si} \right) \left(\sum_{t,u \in B} \alpha_{rs}^{tu} \sum_{i \in N} c_{ti} c_{ui} + \beta_{rs} \right) + \gamma,$$

so that the objective function of the closed-shell formulation is:

$$\min_{c \in \mathbb{R}^{bn}} E^{\text{closed}}(c),$$

where $b = |B|$. We remark that $E^{\text{closed}}(c)$ refers to the energy of closed-shell systems assuming that $\{\varphi_i \mid i \in N\}$ is an orthonormal set. We therefore impose orthonormality constraints:

$$\forall i \le j \in N \quad \langle \varphi_i | \varphi_j \rangle = \delta_{ij},$$

where δ_{ij} is the function with value 1 if and only if $i = j$. Replacing $\varphi_i = \sum_{r \in B} c_{ri} \chi_r$ for all $i \in N$ and $\langle \chi_r | \chi_s \rangle = S_{rs}$ for all $r, s \in B$, we obtain a constraint expressed in function of the decision variables c:

$$\forall i \le j \in N \quad \sum_{r,s \le b} S_{rs} c_{ri} c_{sj} = \delta_{ij}.$$

We remark that $S_{rr} = 1$ for all $r \in B$. Further constraints to the problem are given by the decision variable bounds:

$$c^L \le c \le c^U,$$

which depend on the instance. Common values for these bounds are ± 1.5 or ± 2. The HFP can thus be formulated as follows [12,18]:

$$\left. \begin{array}{r} \min\limits_{c \in \mathbb{R}^{bn}} \quad E^{\text{closed}}(c) \\ \forall i \le j \in N \sum\limits_{r,s \le b} S_{rs} c_{ri} c_{sj} = \delta_{ij} \\ c^L \le c \le c^U, \end{array} \right\} \tag{3}$$

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

3 Formulation for open-shell systems

Orthonormality of the $\{\varphi_i\}$ set derives from physical considerations in closed-shell systems, but is only imposed to simplify SCF computations in open-shell systems [19,8]. In such systems, $\{\varphi_i\}$ is simply required to be a linearly independent set. The molecular energy expression changes accordingly [3]:

$$E^{\text{open}}(c) = \sum_{\substack{i \in N \\ r,s \in B}} W_i \beta_{rs} c_{ri} c_{si} + \sum_{\substack{i,j \in N \\ r,s,t,u \in B}} A_{ij} \alpha_{rs}^{tu} c_{ri} c_{si} c_{tj} c_{uj} - \sum_{\substack{i,j \in N \\ r,s,t,u \in B}} D_{ij} \alpha_{rs}^{tu} c_{ri} c_{sj} c_{tj} c_{ui}, \qquad (4)$$

where W is a given constant n-vector and A, D are given constant $n \times n$ matrices, where $n = |N|$. Moreover, since we no longer impose the orthonormality of $\{\varphi_i\}$, we need to explicitly require that each φ_i be normalized:

$$\forall i \in N \quad \sum_{r \in B} S_{rr} c_{ri}^2 = 1.$$

Modelling linear independence of a general set of vectors $Y = \{\varphi_i \in \mathbb{R}^b \mid i \in N\}$, where $b = |B|$, within a MP formulation is far from easy. Specifically, the direct formulation:

$$\forall \lambda \in \mathbb{R}^n \quad \sum_{i \in N} \lambda_i \varphi_i \neq 0 \qquad (5)$$

is an uncountable set of constraints, each of which defines an open set in $\mathbb{R}^b$, i.e. the entire space minus a proper affine subspace. The uncountable intersection of open sets is certainly not guaranteed to be closed, which means that not only the corresponding MP formulation would have a practically unmanageable number of constraints, but also that no minimum might exist within the feasible set described by it. It is relatively easy to find examples that show that replacing (5) with any finite subset of its constraints fails to guarantee linear independence. An altogether different idea stems from the following equivalent condition to the linear independence of Y: the Cayley-Menger determinant [2] relative to Y in $\mathbb{R}^n$ has nonzero value if and only if Y is linearly independent. This is equivalent to requiring that the $(n-1)$-dimensional simplex defined by the projection of Y onto $\mathbb{R}^{n-1}$ should have positive volume. Since explicit formulæ for computing an $n \times n$ determinant are only valid for a given constant n, or else are recursive, it is not clear how to express this condition in practice in a MP formulation in such a way as to be valid for any n. Moreover, even if a valid expression were found, in general it would yield a polynomial MP of degree n, which is considered to be very hard to solve for even moderately high values of n.

In general, the linear independence constraint is a special case of the more general "rank constraint", i.e. requiring that a matrix of decision variables should have a given rank r. Although a famous relaxation of this constraint yields the well-known Semidefinite Programming (SDP) problem [22], the paper [15] is, to the best of our knowledge, the only existing reference treating this problem explicitly. The formulation given in [15] is a non-convex quadratic MP formulation whose constraints require the projection of Y on $\mathbb{R}^n$ to yield a square symmetric matrix with n nonzero eigenvalues. In the present setting, after

replacing φ_i with the corresponding expression in terms of the c (see Eq. (2)), we obtain the following eigensystem:

$$\forall k, i \in N \quad \sum_{\substack{j \in N \\ r,s \in B}} S_{rs} c_{ri} c_{sj} v_{kj} = \lambda_k v_{ki}, \tag{6}$$

where $\{\lambda_k \in \mathbb{R} \mid k \in N\}$ is the set of eigenvalues, and $\{v_k \in \mathbb{R}^n \mid k \in N\}$ is the set of corresponding eigenvectors. We remark that since the matrix is expressed in terms of decision variables c, both λ and v are also decision variables of the new MP formulation. Without loss of generality, we require the eigenvectors to be an orthonormal set:

$$\forall k \leq h \in N \quad \sum_{j \in N} v_{kj} v_{hj} = \delta_{kh}.$$

Finally, we introduce a constant $\varepsilon > 0$ and require that all squared eigenvalues should have greater value:

$$\forall k \in N \quad \lambda_k^2 \geq \varepsilon.$$

This ensures all eigenvalues are nonzero, which is equivalent to Y being a linearly independent set.

In summary, the following:

$$\left. \begin{array}{rl}
\min\limits_{c,\lambda,v} & E^{\mathsf{open}}(c) \\
\forall i \in N & \sum\limits_{r \in B} S_{rr} c_{ri}^2 = 1 \\
\forall k, i \in N & \sum\limits_{\substack{j \in N \\ r,s \in B}} S_{rs} c_{ri} c_{sj} v_{kj} = \lambda_k v_{ki} \\
\forall k \leq h \in N & \sum\limits_{j \in N} v_{kj} v_{hj} = \delta_{kh} \\
\forall k \in N & \lambda_k^2 \geq \varepsilon \\
& c^L \leq c \leq c^U
\end{array} \right\} \tag{7}$$

is a valid formulation for the HFP on open-shell systems for any $\varepsilon > 0$. It consists of the minimization of a quartic polynomial subject to a set of cubic and quadratic polynomials. We remark that since orthonormal vectors are also linearly independent normalized vectors, the feasible set of (7) contains the feasible set of (3).

4 Reformulations and solution methods

Both formulations (3) and (7) belong to the class of nonconvex NLPs: they are multi-extremal, involve nonlinear terms in both the objective function and the constraints, and are defined over continuous variables only. This calls for GO solution methods. We had previously [11, 12, 18] employed spatial Branch-and-Bound (sBB) algorithms in order to solve the small closed-shell systems He and Be: the advantage of the sBB is that it provides a guarantee of ε-approximation for any given $\varepsilon > 0$. For the larger atomic systems tested in this paper, however, the sBB is too slow.

Some computational experience on HFP instances showed that the closed-shell MP formulation (3) is numerically rather well-behaved notwithstanding its nonconvexities: a few well-known local NLP solvers, such as SNOPT [7] and IPOPT [4], usually identify the global optima if restarted from a few different starting points. This suggests that a simple MS metaheuristic [23] is appropriate to solve (3).

The open-shell MP formulation (7), however, poses severe numerical problems to all local NLP solvers we tried (SNOPT, IPOPT, Filter [5], MINOS [20] and DONLP2 [26]). A little testing showed that solvers had a hard time trying to satisfy (6). We therefore reformulated (7) by adding slack variables s_{ki} (for $k, i \in N$) to (6) whilst minimizing their squares added to the objective (4). This results in the following formulation:

$$
\left.
\begin{aligned}
\min_{c, \lambda, v, s} \quad & E^{\text{open}}(c) + \sum_{k,i \in N} s_{ki}^2 \\
\forall i \in N \quad & \sum_{r \in B} S_{rr} c_{ri}^2 = 1 \\
\forall k, i \in N \quad & \sum_{\substack{j \in N \\ r,s \in B}} S_{rs} c_{ri} c_{sj} v_{kj} = \lambda_k v_{ki} + s_{ki} \\
\forall k \le h \in N \quad & \sum_{j \in N} v_{kj} v_{hj} = \delta_{kh} \\
\forall k \in N \quad & \lambda_k^2 \ge \varepsilon \\
& c^L \le c \le c^U,
\end{aligned}
\right\}
\tag{8}
$$

whose numerical behaviour with respect to local NLP solvers is improved. Naturally, unless the computed optimum $(c^*, \lambda^*, v^*, s^*)$ of (8) has $\sum_{k,i \in N} s_{ki}^* = 0$, we cannot be sure that linear independence is attained. In practice, however, this is not a problem, as linear independence of the solution can be verified *a posteriori*. Should the solution fail to be linearly independent, a new run of the (stochastic) MS algorithm can be launched. Because the MS converges to the global optimum with probability 1 in infinite time [25], the chance that all computed optima turn out to be linearly dependent is low.

It is interesting to remark that (8) exhibits some symmetry: with the computational methods discussed in [16] and the generalization from instance to problem proposed in [17], the formulation group of (8) is found to be the full symmetric permutation group S_n (where $n = |N|$), and corresponds to the fact that no constraint in (8) enforces the eigenvalues λ_k to be ordered. Adjoining constraints

$$
\lambda_1 \le \ldots \le \lambda_n
\tag{9}
$$

to (8) to the formulation yields a reformulation of the narrowing type (see [14]) with trivial group. This is generally useful when the solution approach is of the sBB type. For metaheuristics, as remarked in [17], it is better to keep formulation symmetry.

5 Computational results

We implemented the formulations (3) and (8) using the MP language AMPL [6] and solved four closed-shell and two open-shell instances using the MS approach with both SNOPT and IPOPT. It is worth remarking that there is nothing complicated about our MS: it simply computes random starting points, then runs the SNOPT local solver (which is, itself, a very sophisticated piece of mathematical software) a finite number I of times (in all our experiments, $I = 10$). We want to convey the message that the *model-and-solve* paradigm we are using is advantageous: a sophisticated model may well drive a general-purpose solver to an early and successful termination, thereby yielding an efficient overall method. The accent is on modelling the problem rather than devising and implementing a solution algorithm.

In Table 1 we recorded (in column order): the atomic symbol, the number V of variables in the formulation, the number C of constraints in the formulation, the value of the optimal energy E^* expressed in atomic units (au), the seconds of user CPU time computed using a single core of an Intel Core i7-7500 CPU running at 2.70GHz with 16GB RAM, and the MS iteration (over $I = 10$) at which the putative global optimum was found.

Atom	Name	V	C	E^* (au)	CPU (s)	Opt. It.
Closed-shell systems						
He	helium	2	1	-2.74706	0.01	8
Be	beryllium	4	3	-14.3519	0.02	8
Ne	neon	25	15	-126.603	0.11	6
Mg	magnesium	54	21	-196.687	3.95	7
Open-shell systems						
Li	lithium	20	11	-12.2584	0.08	9
B	boron	36	23	-57.334	0.17	4

Table 1 Instances at `www.lix.polytechnique.fr/~liberti/hartreefock-instances.zip`.

The energy values E^* are not as precise as those given in the literature. This is a consequence of two factors. First, the data we used (in particular the values for the S_{rs} symbols) was obtained through a simplified model. Second, the MP formulations we employ allow a 1×10^{-3} tolerance in order to improve the numerical stability of the associated NLP solvers. Because the energy functions E^{closed} and E^{open} depend quartically on the decision variables, the imprecision in the objectives generally corresponds to smaller imprecisions in the values of the c variables. This need not be a limitation: the solutions we obtain are intended to provide very high quality starting points for further refinement using the established SCF methodology. In order to validate our numerical results, the MP formulations were solved by MS using very different local solvers, namely SNOPT (our main solver, to which the CPU times refer), IPOPT, KNitro, Filter, MINOS, Loqo. The (putative) global optima were confirmed by all solvers (except for some nonconvergence failures of MINOS on Ne and Mg). In Table 1, however, we only reported results relative to the SNOPT solver, which we found to be the most efficient (with Knitro and Loqo) for this data set.

6 Conclusion

In this paper, we extended the Mathematical Programming approach for solving the Hartree-Fock equations to open-shell systems. This methodology was first proposed in [12] limited to closed-shell systems. The present extension is innovative as far as quantum chemistry is concerned, because it dispenses with a limitation imposed by numerical requirements of the SCF methodology, i.e. orthogonality of the $\varphi_i's$. As far as mathematical programming is concerned, the approach is innovative because insofar as it discusses the "linear independence constraint", which is sometimes useful but not easy to encode in MP terms.

Acknowledgments

References

1. Belotti, P., Lee, J., Liberti, L., Margot, F., Wächter, A.: Branching and bounds tightening techniques for non-convex MINLP. Optimization Methods and Software **24**(4), 597–634 (2009)
2. Blumenthal, L.: Theory and Applications of Distance Geometry. Oxford University Press, Oxford (1953)
3. Cardozo, T., Chaer Nascimento, M.: Private communication (2012)
4. COIN-OR: Introduction to IPOPT: A tutorial for downloading, installing, and using IPOPT (2006)
5. Fletcher, R., Leyffer, S.: User manual for FILTER. Tech. rep., University of Dundee, UK (1999)
6. Fourer, R., Gay, D.: The AMPL Book. Duxbury Press, Pacific Grove (2002)
7. Gill, P.: User's Guide for SNOPT 5.3. Systems Optimization Laboratory, Department of EESOR, Stanford University, California (1999)
8. Hollauer, E., Nascimento, M.: A generalized multistructural wave function. Journal of Chemical Physics **99**, 1207 (1993)
9. Janes, P.: Rigorous numerical approaches in electronic structure theory. Ph.D. thesis, Australian National University (in revision)
10. Janes, P., Rendell, A.: Deterministic global optimization in ab-initio quantum chemistry. Journal of Global Optimization **56**(2), 537–558 (2013)
11. Lavor, C., Cardozo, T., Nascimento, M.A.C.: Using an interval branch-and-bound algorithm in the Hartree-Fock method. International Journal of Quantum Chemistry **103**, 500–504 (2005)
12. Lavor, C., Liberti, L., Maculan, N., Chaer Nascimento, M.: Solving Hartree-Fock systems with global optimization methods. Europhysics Letters **5**(77), 50,006p1–50,006p5 (2007)
13. Levine, I.: Quantum Chemistry, second edn. Prentice-Hall, Upper Saddle River, New Jersey (2000)
14. Liberti, L.: Reformulations in mathematical programming: Definitions and systematics. RAIRO-RO **43**(1), 55–86 (2009)
15. Liberti, L.: Modelling rank constraints in mathematical programming. In: S. Pickl *et al.* (eds.) Proceedings of the 11th Cologne-Twente Workshop on Graphs and Combinatorial Optimization. Universität der Bundeswehr, München (2012)
16. Liberti, L.: Reformulations in mathematical programming: Automatic symmetry detection and exploitation. Mathematical Programming A **131**, 273–304 (2012)
17. Liberti, L.: Symmetry in mathematical programming. In: J. Lee, S. Leyffer (eds.) Mixed Integer Nonlinear Programming, *IMA*, vol. 154, pp. 263–286. Springer, New York (2012)
18. Liberti, L., Lavor, C., Nascimento, M.C., Maculan, N.: Reformulation in mathematical programming: an application to quantum chemistry. Discrete Applied Mathematics **157**, 1309–1318 (2009)
19. Martin, R.: Electronic Structure: Basic Theory and Practical Methods. Cambridge University Press, Cambridge (2004)
20. Murray, W., Gill, P., Raman, R., Kalvelagen, E.: GAMS/MINOS: A solver for large-scale nonlinear optimization problems. Tech. Rep. 10.1.1.16.8658, CiteSeerX (2007)
21. Penrose, R.: The road to reality. Knopf (2004)
22. Rendl, F.: Semidefinite relaxations for partitioning, assignment and ordering problems. 4OR **10**(4), 321–346 (2012)
23. Schoen, F.: Two-phase methods for global optimization. In: P. Pardalos, H. Romeijn (eds.) Handbook of Global Optimization, vol. 2, pp. 151–177. Kluwer Academic Publishers, Dordrecht (2002)
24. Schrödinger, E.: Quantisierung als Eigenwertproblem. Annalen der Physik **384**:361–377 (1926)
25. Solis, F., Wets, R.: Minimization of random search techniques. Mathematics of Operations Research **6**(1), 1981 (19-30)
26. Spellucci, P.: DONLP2 Users Guide. TU Darmstadt (2009)
27. Szabo, A., Ostlund, N.: Modern Quantum Chemistry. McGraw-Hill, New York (1989)
28. Zorn, K., Sahinidis, N.: Hartree fock self-consistent calculations: Global optimization of electronic structure. In: Proceedings of 10th AIChE Annual Meeting. AIChE (2010)